Università degli Studi di Napoli "Federico II"

Dipartimento di Scienze Fisiche

Dottorato di Ricerca in Fisica Fondamentale ed Applicata
XXIV ciclo (2008-2011)

Thesis for Doctor of Philosophy degree
submitted by **Elisa Musto**

The search for the rare decay $B_s \to \mu^+\mu^-$

with the ATLAS experiment at LHC

Advisors :

Prof.ssa Mariagrazia Alviggi
Dott. Domenico Della Volpe

Coordinator:

Prof. Raffaele Velotta

Contents

Introduction

ATLAS (***A*** ***T****oroidal* ***L****HC* ***A****pparatu****S***) is a general-purpose high-energy physics experiment built to exploit at best the Physics potential offered by the proton-proton ***L****arge* ***H****adron* ***C****ollider* (LHC) at CERN. The LHC started its operation in late 2009 reaching a center of mass energy of 7 TeV and delivering up to now an integrated of luminosity about 5.6 fb^{-1}, with a peak instantaneous luminosity of $\sim 3.5\times10^{33}$ $\text{cm}^{-2}\text{s}^{-1}$.

The LHC allows to study physics processes to an unexplored energy scale ($\sim$ 1 TeV) and, at its nominal instantaneous luminosity of 10^{34} $\text{cm}^{-2}\text{s}^{-1}$, it will be possible to measure physics processes with cross sections smaller than 10^{-6} nb.

Among the main goals of the LHC research program, there is the search for the Higgs boson and for New Physics Beyond the Standard Model; to the latter extent the $B_s \to \mu^+\mu^-$ decay, object of this thesis work, is considered a golden channel.
It is a *Flavor Changing Neutral Currents* process, where the direct transition from $b \to s$ quark in the Standard Model is suppressed and occurs at the lowest order through one-loop diagrams; in this picture, the theoretical estimation of the branching ratio for the $B_s \to \mu^+\mu^-$ decay is $\sim 3.2\times10^{-9}$. New Physics processes can significantly enhance it, so its measurement can be a sensitive tools to detect them.

The strategy adopted for this data analysis is the so called "blind analysis" technique, i.e. the signal is only modeled by a Monte Carlo, while the data are used to model the background after removing ("blinding") all the events with a di-muon invariant mass in a selected region around the B_s mass peak. All the study, the evaluation of the errors and evaluation of the quantities involved, are done without looking at the data in the blinded region. Only once the analysis has been approved by the ATLAS community, the data are unblinded and the result obtained made public.

The analysis here presented is based on the data collected in 2010 data taking period, which correspond to a integrated luminosity of $\simeq \ 40pb^{-1}$. This statistics results very limited, expecially if compared to the $\ 5.6fb^{-1}$ of

the 2011, but the analysis based on the 2011 data, to which I am contributing, is still blind.

Indeed, most of the technique developed, tested and validated here are the baseline of 2011 data analysis, and as recognition of this work, the ATLAS physics coordination has allowed me to unblind the 2010 data and produce a result, even if the official result of the Collaboration will be based maily on 2011 data.

The quantity to measure is the $B_s \to \mu^+\mu^-$ branching fraction which, in order to reduce systematic uncertainties, is leaded back to the study of a relative branching ratio with respect to a well-measured reference signal B_{ref}:

$$Br(B_s \to \mu\mu) = BR(B_{ref}) \times \frac{\mathcal{L}_{ref}}{\mathcal{L}_{\mu\mu}} \times \frac{\epsilon_{ref}\alpha_{ref}}{\epsilon_{\mu\mu}\alpha_{\mu\mu}} \times \frac{f_{ref}}{f_s} \times \frac{N_{\mu\mu}}{N_{ref}}$$

where, for each decay mode, N is the number of observed events, ϵ and α are, respectively, the efficiencies and acceptances, and f_{ref}/f_s is the relative $p-p$ production rate of B_{ref} and B_s mesons. The main sources of systematic uncertainties ($\mathcal{O}$ 10 %) are given by the number of signal events $N_{\mu\mu}$ and by the cross sections ratio f_{ref}/f_s . Often the number of signal events is expressed in terms of the Single Event Sensitivity (SES); this quantity is very important as it gives an indication of the experiment sensitivity to this search, containing all the elements to determine it but being independent on the number of event (i.e. the integrated luminosity) and is related to the branching ratio through a multiplicative factor:

$$Br(B_s \to \mu\mu) = SES \times N_{\mu\mu}$$

Given its low probability of being observed, instead of a real measurement of the $Br(B_s \to \mu\mu)$ an Upper Limit on its value, with a confidence level, is set. This limit will depend on the number of observed event after a selection procedure, compared with what are the expectation coming from the Monte Carlo simulation.

A first step of the study has been the estimation of the reference channel yield, N_{ref}, in ATLAS data by using maximum likelihood fitting techniques. To deal with the systematic uncertainties deriving from the non-correct description of the B mesons kinematics and of the detector response in Monte Carlo, a data-driven iterative procedure has been set-up in order to re-weight these models according to data. A "cut and count" approach has been used to separate signal from background events. This approach has been optimized in order to maximize the analysis sensitivity by using the Punzi estimator [84], which has been proven to be very effective in this kind of search.

After the optimization, the cuts are applied to the data and, using the two sidebands around the blinded region, the non resonant background has

been modeled and used to extrapolate the number of expected background events in the signal region.

Finally, the signal region has been unblinded: the number of observed events, together with all other quantities and the corresponding uncertainties already determined, has been used to set the Upper Limit. The latter has been estimated to be $Br(B_s \to \mu^+\mu^-) < 5.91\times 10^{-7}$ at 95% CL. This is not yet competitive with the last value quoted by other experiments which in turn used a significantly higher statistics (CMS, CDF) or, as LHCb, have an higher sensitivity, being designed only for B- Physics measurements, but it proves that the ATLAS sensitivity is good enough to allow to produce a competitive result with the much higher statistic of the 2011 data.

This thesis work is divided into five chapters. In Chapter 1 the Physics at LHC and his physics goals are illustrates, like the search for the Higgs bosons and the search for New Physics.

In Chapter 2 the focus is on the B mesons physics; in particular, *Flavor Sector* and *Flavor Changing Neutral Currents* are described in the Standard Model; then the expression for $B_s \to \mu^+\mu^-$ amplitude in the effective theory picture of the *Operator Group Expansion* is shown. Finally, an overview of the theories Beyond Standard Model that may give deviation with respect to the Standard Model expectations for the decay amplitude under study is given.

Chapter 3 is dedicated to the description of the experimental apparatus: the ATLAS detectors and the Computing Model are described and an overview of the performance so far achieved is given.

The last two chapters are dedicated to a detailed review of the $B_s \to \mu^+\mu^-$ search with the ATLAS detector. In Chapter 4 all the quantities entering in the analysis are enlightened, while in Chapter 5 the ATLAS result for the Upper Limit on the branching ratio and the comparison with the other recent experimental results are presented.

Chapter 1

LHC Physics

The LHC features render this machine the best candidate to explore high energy phenomena that still need to be understood, like the Electroweak Symmetry Breaking or the Physics beyond the Standard Model energy scale. In this chapter the proton-proton collision phenomenology is described, giving emphasis to the expectations and the results in the various research fields explored by the collider.

1.1 Proton-proton collisions at LHC

The Large Hadron Collider (LHC) [1] is a superconducting proton-proton collider housed in the $27Km$ long tunnel that previously hosted the Large Electron Positron (LEP) collider at CERN.

LHC has been designed to reach an instantaneous luminosity of 10^{34} cm^{-2} s^{-1} at a center of mass energy of $\sqrt{s} = 14$ TeV in $p - p$ collisions, but it will also collide heavy nuclei (Pb-Pb) with a center of mass energy of 1150 TeV, corresponding to 2.76 TeV u^{-1} and 7 TeV per charge, and a peak luminosity of 10^{27} cm^{-2} s^{-1}.

Four experiments are located in the four interaction points of the beams along this tunnel (see fig 1.1).

Two of them, ATLAS (A *Toroidal LHC ApparatuS*) [2] and CMS (*Compact Muon Solenoid*) [3], have been designed to work at high luminosities in order to allow them to perform a large variety of physics studies.

The other two, instead, are special purpose experiments, which work at lower luminosities ($\sim 10^{30}$) cm^{-2} s^{-1} in order to optimize the number of collisions to perform dedicated studies. In particular, LHCb (Large Hadron Collider *beauty*) [4] is dedicated to investigate the b-quark properties and the CP-violating decays, while ALICE (A *Large Ion Collider Experiment*) [5] was

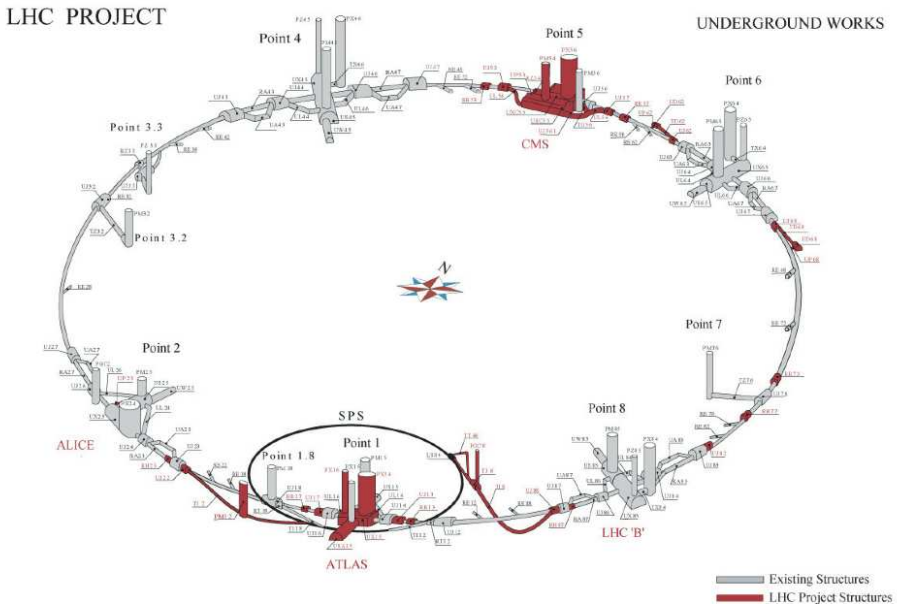


Figure 1.1: The LHC ring.

built to explore the behavior of strongly interacting particles at high density and temperature, emulating the early stage of the Universe by colliding heavy ions.

Finally, two forward detectors, LHCf [6] and TOTEM [7], have been designed in order to study physics processes in the region very close to the particles beam, thus complementing the LHC general-purpose experiments physics programs.

The foreseen start-up of data taking has been delayed because of an accident occurred in September 2008 during the initial commissioning [8]. As a consequence of this incident it was decided to run at a center of mass energy of $\sqrt{s} = 7$ TeV until the next upgrade of the machine that will allow to reach higher energies; LHC restarted its operations in November 2009 with two circulating beams at a center-of-mass energy of 900 GeV, and soon after the center-of-mass energy was raised up to 2.36 TeV, thus exceeding the previous energy record held by Tevatron.
The first p-p collisions at $\sqrt{s}$ =7 TeV took place on 30th March 2010.

Even if far below the nominal ones, such high energies have never been reached before by any other collision machine, as testified by figure 1.2 in which the LHC stored energy and the ones stored by the previous experiments are compared. The two beams of protons reach an energy of 3.5 GeV in different steps through a complex acceleration system

Beam parameters	
Beam energy	7 TeV
Injection energy	450 GeV
Maximum luminosity	$10^{34}\text{cm}^{-2}\text{s}^{-1}$
Time between collisions	25 ns
Bunch length	7.7 cm
RMS beam radius at the interaction point	16.7 μm
Technical parameters	
Ring length	26658.9 m
Radiofrequency	400.8 MHz
Number of bunches	2808
Number of protons per bunch	$1.15{\cdot}10^{11}$
Number of dipoles	1232
Dipole magnetic field	8.33 T

Table 1.1: The LHC operational parameters for p-p collisions.

- they are first accelerated by a linear accelerator (LINAC2) to 50 MeV;
- the Proton Synchrotron Booster (PSB) then increases their energy up to 1.4 GeV;
- they are further injected into the Proton Synchrotron (PS) were they are accelerated to 25 GeV and collected in bunches;
- the bunches pass then into the Super Proton Synchrotron (SPS) where they are accelerated to 450 GeV;

Operational parameter	2010	2011	Nominal
Energy $[TeV]$	3.5	3.5	7
β^* $[m]$	3.5,	0.55	1.0
Emittance $[\mu m]$	2.5	2	3.75
Number of bunches	368	1404	2808
Bunch population $[\#ofprotons]$	$1.2{\times}10^{11}$	$1.6{\times}10^{11}$	$1.15{\times}10^{11}$
Stored energy $[MJ]$	28	100	360
Peak luminosity $[cm^{-2}s^{-1}]$	$2{\times}10^{32}$	$3.5{\times}10^{33}$	$1{\times}10^{34}$

Table 1.2: LHC operational parameters: comparison between 2010 and 2011 LHC achievements in p-p collisions and nominal reach.

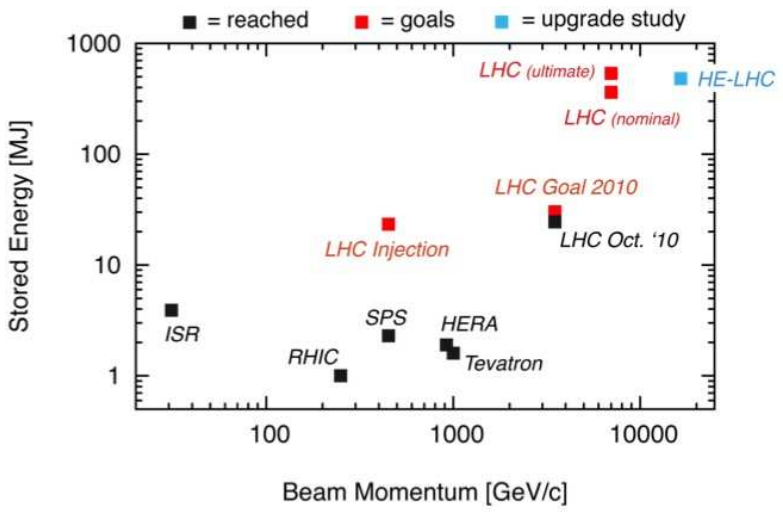


Figure 1.2: Comparison between the LHC stored energy and the other colliders.

- finally they are injected into the main ring, the LHC, where they are accelerated to the energy at which collisions happen.

The beam has a bunch structure, and bunches are spaced by $\sim 25ns$. This structure has been adopted in order to maximize the luminosity, have a number of interactions per crossing within certain limits and keeping track of the beam parameters. This structure has been taken into account also in the design of the experiment Trigger Systems, which at each bunch crossing are called to decide whether an event is of interest or not.

Into the LHC tunnel, in order to compensate the energy losses due to radiation, super-conducting cavities with a field gradient of 5 MV/m are installed. Super-conducting dipole magnets used to bend the protons have been designed to produce a magnetic field of up to ∼8 T at nominal energies; several hundred magnets are further used to focus the beams.

A summary of the main LHC nominal operational parameters can be found in table 1.1; moreover, table 1.2 reports the comparison between some of the main LHC nominal parameters and the values achieved in 2010 and 2011.

The total LHC proton-proton cross-section σ, together with his main components, is shown in figure 1.3 as a function of the center of mass energy; at LHC energies it is of the order of 100 mb.

The cross section is related to the luminosity $\mathcal{L}$ of the machine and to the expected rate R of events produced by a p-p interaction per second through

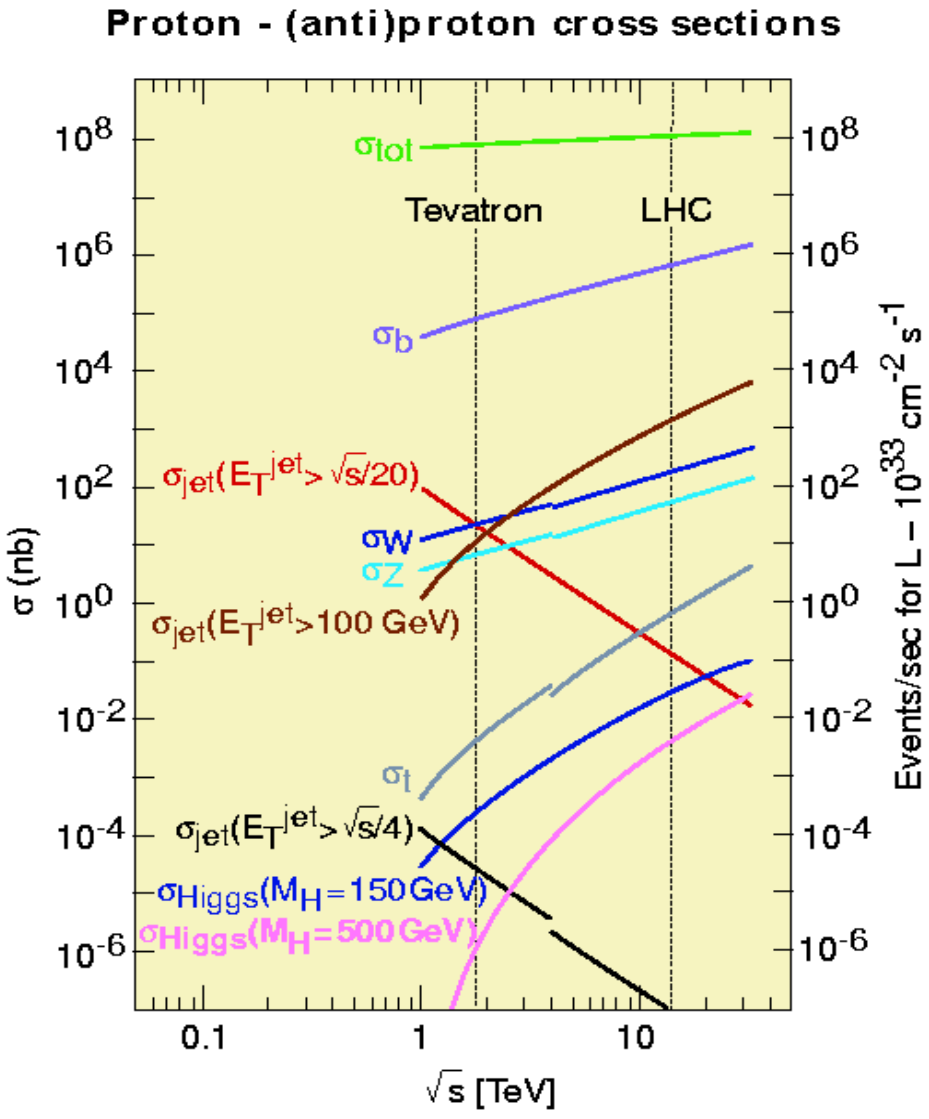
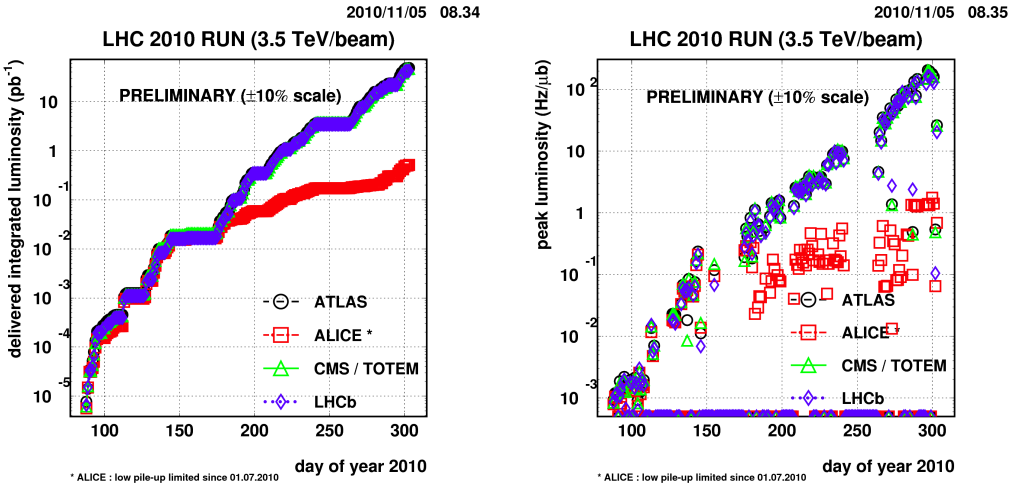
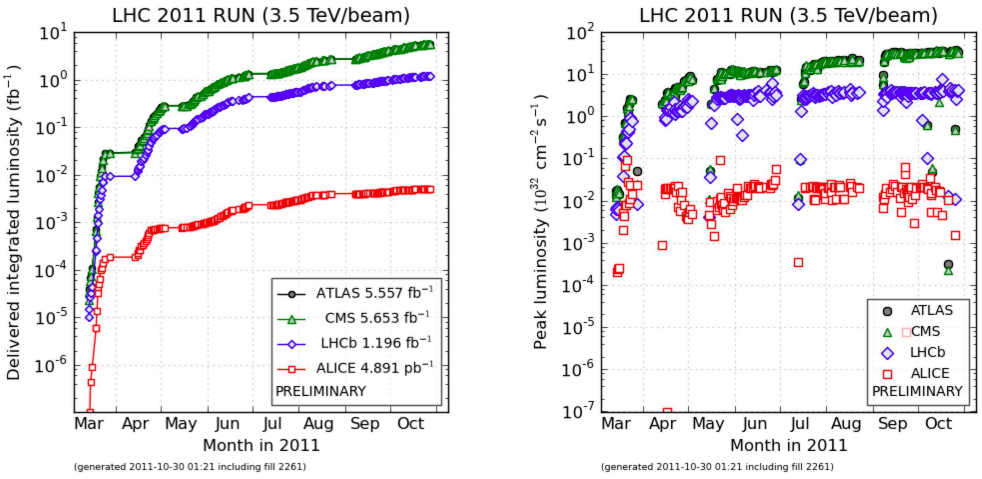


Figure 1.3: LHC total cross section.



(a) 2010 p-p Integrated luminosity (b) 2010 p-p peak luminosity

Figure 1.4: Proton-proton collisions: integrated (a) and instantaneous (b) luminosity in 2010.



(a) 2011 p-p Integrated luminosity (b) 2011 p-p peak luminosity

Figure 1.5: Proton-proton collisions: integrated (a) and instantaneous (b) luminosity in 2011.

the equation

$$R = \sigma \times \mathcal{L}. \tag{1.1}$$

The instantaneous luminosity depends only on the beam parameters [9], as can be read by the expression

$$\mathcal{L} = \frac{N_b^2 n_b f_r F}{4\pi\sigma_x^*\sigma_y^*}, \tag{1.2}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f_r the revolution frequency, $4\pi\sigma_x^*\sigma_y^*$ is the transverse beam section [1] assuming a gaussian beam profile, and F is a geometric reduction factor that takes into account the crossing angle between the beams at the Interaction Point (IP) [2].

The instantaneous and integrated luminosity reached by LHC are illustrated in figures 1.4 and 1.5, respectively for 2010 and 2011 data-taking [10]: the first year allowed to collect 45 pb^{-1} of integrated luminosity in p-p collisions, while, thanks to the increasing of instantaneous luminosity in 2011, with an equivalent data-taking time, 5.2 fb^{-1} of integrated luminosity was delivered to the experiments.
The increasing of luminosity has been reached in a first step by increasing the number of protons per bunch N_b, as equation 1.2 depends on the square of this quantity; nevertheless this parameter could not be increased indefinitely since the more are the protons inside a bunch the more the events result complicate to analyze because of the multiple interactions. A further increasing of the luminosity has been reached by "squeezing" the beams, reducing the beam spreads.

By inserting the nominal peak luminosity in equation 1.1 one gets for the rate:

$$R = 100\,mb \times 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1} \simeq 10^{9}\mathrm{Hz} \tag{1.3}$$

Even by using the current peak luminosity of the machine, that is about one third of the nominal one, this corresponds to an enormous amount of data that cannot be stored by the various experiments, so a reduction of the rate is

[1] The beam section can be expresses also as $\sigma_x^*\sigma_y^* = \epsilon_n\beta^*/\gamma$, where ϵ_n is the normalized transverse emittance, γ is the relativistic γ function and β^* the relativistic β function at the Interaction Point.

[2] This reduction factor reads $F = 1/\sqrt{\frac{\theta_c\sigma_z}{2\sigma^*}}$, where θ_c is the crossing angle between the beams ($\theta_c \sim 300\mu rad$), σ_z the root mean square of the bunch length and σ^* the transverse beam size at the IP. The crossing angle, although involves a reduction of the luminosity of $\approx 17\%$, is necessary to reduce beam-beam interactions that are caused by the force due to the electromagnetic field of one beam on the particles in the other beam.

necessary; this reduction up to $\sim$ 300 Hz is performed on line by the various experiments through the Trigger System.

In particular, the interesting events can be primarily classified into *soft* and *hard* collisions.

First category groups collisions in which the impact parameter between the incident partons is high, thus the momentum transferred to the out-coming particles is low and prevalently aligned with the beam axis direction since these particles are scattered with small angles with respect to this axis. The transverse momentum (p_T) is typically of the order of $\simeq$ 500MeV, and the events presenting these features are usually denoted as "minimum bias" since they correspond to the minimal request that a trigger system may have, i.e. a collision event.

Hard collision events, instead, are the interesting ones; they correspond to the head-on collisions between the incoming partons and an high transverse momentum is transferred to the out-coming particles. The latter may be scattered at high angles with respect to the beam axis or massive particles in the final states may appear.

It is relevant to note that the center of mass energy of the colliding partons $\sqrt{\hat{s}}$ is different from the center of mass energy of the protons $\sqrt{s}$, but is related to the latter through the relationship:

$$\sqrt{\hat{s}} = \sqrt{x_a x_b s} \tag{1.4}$$

where x_a and x_b are transverse momentum fractions carried by the colliding partons. When approximating $x_a \simeq x_b \equiv x$, one has $\sqrt{\hat{s}} \simeq x\sqrt{s}$; this means that even when the carried fraction of the transverse momentum is of the order of a few %, for $\sqrt{s}$ = 7 TeV a $\sim$ 100 GeV massive particle can be produced.

The advantage of reaching such high energy in the collisions has, as a counterpart, a price to be paid; the experimental environment is, in fact, much less clean than the ones in the lepton colliders and interesting events should be selected among a huge QCD background, which dominates the total cross section, the underlying event [3] and the multiple-parton interactions[4] [11].

By using the nominal rate obtained in equation 1.3, which corresponds to a bunch collision every t_b = 25 ns, it is possible to compute the mean number of simultaneous soft interactions at the nominal luminosity the equation

$$\langle n_c \rangle = R \times t_b \simeq 25. \tag{1.5}$$

[3]The underlying event denotes all event activity except the high-pt jets from the hard scattering process.

[4]When hadrons collide at high-energy, each of the incoming hadrons may be viewed as a beam of partons, hence there is a possibility of having several parton-parton interactions.

Thus, in each bunch crossing the interesting event is readout entangled with a large number of minimum bias events, which constitutes the main component of the so-called *pile-up*.

The difficulty in disentangling the various events imposes a fast detector response ($< 25ns$) and a large detector segmentation, while the high flux of particles produced requires an high resistance to the radiation ($\sim 10^{17} neutrons/cm^2$).

1.2 The LHC research programme

Among the others, the LHC research programme includes as primary goals

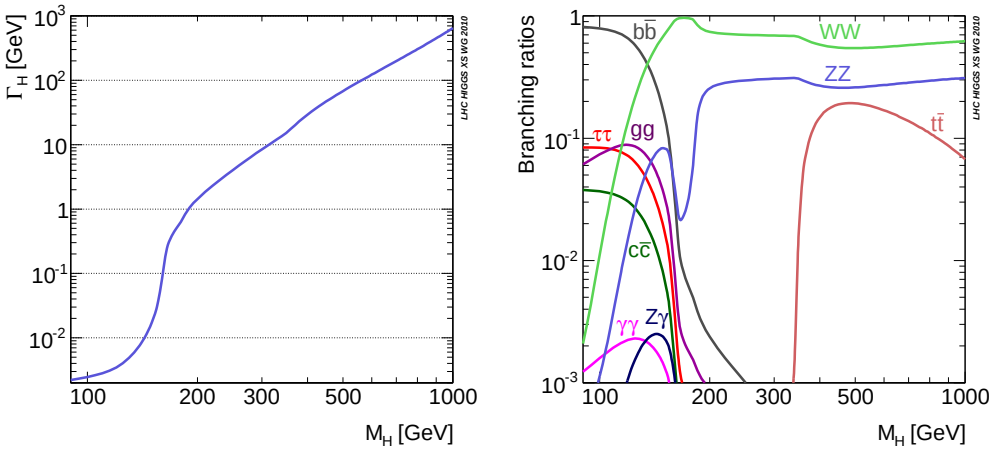
- the search for the Higgs boson;
- detailed studies on physics processes at the Standard Model (SM) energies like
  1. the *top* quark characterization through mass, cross section, polarization and rare decays measurements;
  2. the B mesons physics and the CP violation in the $B^0 - \bar{B}^0$ system, the hunting for *Flavor Changing Neutral Currents* rare decays like $B^0_{d,s} \to \mu^+\mu^-$, $B^0_d \to K^{*0}\mu^+\mu^-$, $B^0_d \to \rho\mu^+\mu^-$ and $B^0_s \to \phi\mu^+\mu^-$;
- the search for New Physics Beyond the Standard Model;
- the heavy-ions collisions studies;

Some of these topics will be covered in the following subsections, and the corresponding latest LHC results will be shown.

1.2.1 The search for the Higgs boson

The Higgs mechanism provides an elegant way of introducing particles masses in the Standard Model (cfr. sec. 2.1) and all the measurement so far performed confirm the predictions with an extraordinary precision.
Nonetheless, even by using the most recent data, the Higgs boson remains unobserved.

One complication in this search is the fact that theory does not predict a value for the Higgs boson's mass but only the relation $m_H = \sqrt{2\lambda}v$, where $v \sim 246$ GeV while λ is an a-dimensional coupling constant and a free parameter



(a) Higgs width (b) Higgs decays

Figure 1.6: (a) the SM Higgs width as a function of the allowed values for the mass. (b) the SM Higgs branching ratios as a function of the allowed values for the mass.

of the theory. The expression for this parameter as a function of the energy scale Q at one-loop [5] reads

$$\lambda(Q^2) = \frac{\lambda(v^2)}{1 - \frac{3}{4\pi^2}\lambda(v^2)\ln\left(\frac{Q^2}{v^2}\right)} \tag{1.6}$$

where v is the electroweak symmetry breaking scale. If the energy is much smaller than the electroweak breaking scale $\lambda \to 0_+$ while in the opposite limit case it can indefinitely grow with energy scale Q, and the boson mass as well.
However to have a consistent and perturbative theory there should be an upper limit on λ, namely $\lambda^* \leq M_P$, where $M_P = G_N^{-1/2} \sim 10^{18}$ GeV is the Planck mass scale at which gravity becomes relevant with respect to the other interactions.

A review of theoretical constraints on Higgs mass can be found in [12]; by using these constraints the upper limit for the mass results ~ 1 TeV, even though at large mass corresponds values for the width of the same magnitude (see left plot in fig. 1.6), thus hardly fitting the concept of a resonance.

The value of the mass also determines the accessible phase-space for the boson decays, the most probably being the ones in particles having high

[5] This expression is the solution of the Renormalization Group Equation for λ: $\frac{d}{dQ^2}\lambda(Q^2) = \frac{3}{4\pi^2}\lambda(Q^2) + \mathcal{O}(Q^4)$.

masses [6]; in this view it is useful to split the mass range into three parts:

1. the "low mass" range, corresponding to $m_H \in [110, 130]$ GeV;
2. the "intermediate mass" range, where $m_H \in [130, 180]$ GeV;
3. the "high mass" range, covering the remaining range, i.e. $m_H \in [180, 1000]$ GeV.

As it is shown in right plot of figure 1.6, in the first mass range the decay $H \to b\bar{b}$ dominates with a branching ratio of $\sim 80\%$; nonetheless, it is very difficult to extract the signal from the huge QCD background, so in this region the searches are performed into less abundant but cleanest decay channels like $H \to \gamma\gamma$ ($\mathcal{O}(10^{-3})$) and $H \to \tau^+\tau^-$ ($\mathcal{O}(8\%)$).

In the intermediate range, instead, the decays $H \to WW^*$ and $H \to ZZ^*$ are preferred; in these decays one of the vector boson is "virtual" because the mass is below the kinematical thresholds for producing both "on shell". The search results easier in the channels where the vector bosons subsequently decay into lepton pairs rather than in hadrons, although the branching ratios for these chains are lower.

The decays into massive bosons dominate also the high mass range, but in this case the Higgs mass allows the production of two real particles. As before, the golden channels are the ones with four leptons in the final states.

Experimental constraints

The previous direct experimental searches for this particle did not have success at both LEP [13] and Tevatron [14], which placed limits on the Higgs mass by excluding, respectively, the ranges $m_H \leq 114.4 GeV$ and $m_H \in [156, 177]$ GeV. ATLAS [15] and CMS [16] published their preliminary results obtained by combining the searches in different decay channels [7] by using about 1-2 fb^{-1} of integrated luminosity.

In both cases no clear evidence of the Higgs boson has been found; in particular, as can be seen by the combined limits illustrated in figures 1.7 and 1.8, at 95% C.L. ATLAS excluded the SM Higgs boson in the mass ranges from 146 GeV to 230 GeV, 256 GeV to 282 GeV and 296 GeV to 459 GeV, while CMS excluded it in two mass ranges, 149-206 and 300-440 GeV, as well as several narrower intervals in between. The limits obtained by combining

[6] This is due to the nature of the Higgs boson interactions whose couplings are proportional to the masses of the particles involved.

[7] Both Collaborations investigated the following decay channels: $H \to \gamma\gamma$, $H \to \tau^+\tau^-$, $H \to b\bar{b}$, $H \to WW^{(*)} \to l\nu l\nu$, $H \to ZZ^{(*)} \to llll(l\nu l\nu)(llqq)$.

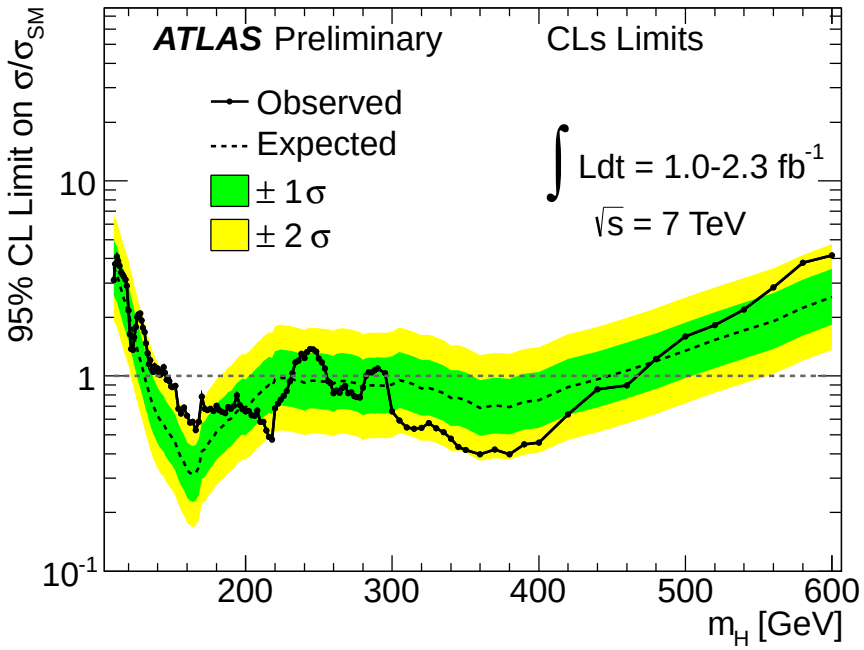


Figure 1.7: ATLAS expected (dashed) and observed (solid) cross section limits at 95% CL, normalized to the Standard Model Higgs boson cross section, as functions of the Higgs boson mass.

ATLAS and CMS results by using $\sim 2fb^{-1}$ of integrated luminosity has been released very recently [17]: as shown in figure 1.9, the SM Higgs boson is excluded at 95% CL or higher in the mass range [141,476] GeV, while the region from 146 to 443 GeV is excluded at the 99% CL, with the exception of three small regions between 220 and 320 GeV.

1.2.2 *Top* Physics

The *top* quark is one of the main attractive SM particles thanks to his peculiar features like

- the value of mass, $m_t \approx 174.5$ GeV, which is of the same order of magnitude of Electroweak Symmetry Breaking scale, $m_{EW} \approx 246$ GeV;
- the value of the lifetime, $\tau_t \approx 0.5{\cdot}10^{-24}$, that is much smaller than the typical time needed for the formation of QCD bound states $\tau_{QCD} \approx 1/\Lambda_{QCD} \approx 3{\cdot}10^{-24}$ s.

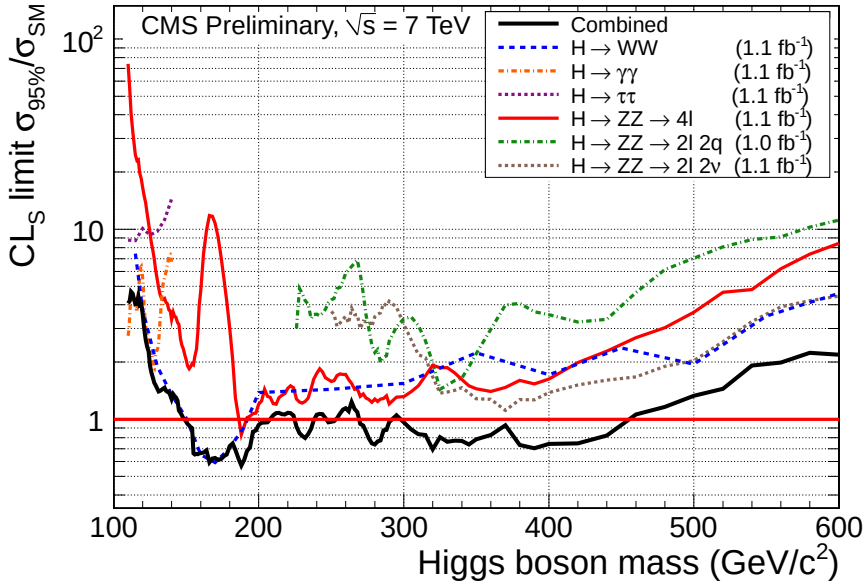


Figure 1.8: CMS expected (dashed) and observed (solid) cross section limits at 95% CL for the individual search channels, normalized to the Standard Model Higgs boson cross section, as functions of the Higgs boson mass.

These features suggest that *top* may have a key role in the electroweak symmetry breaking mechanism: thanks to its heavy mass, in fact, it has a strong coupling to the Higgs boson. Moreover, thanks to its lifetime it decays (predominantly in the Wb channel) before undergoing hadronization, thus offers the opportunity to study the properties of a bare quark.

Besides the SM physics, *top* quark is interesting also because its production is often the largest background to many possible New Physics signals at the LHC.
In particular, its semileptonic decays (corresponding to the ones in which the W boson further decays into a lepton-neutrino pair) produce large missing transverse energy, high energy leptons and hadronic jets, that are the same signatures of many New Physics models like Supersymmetry or Universal Extra Dimensions.

Thanks to its large production cross section, at LHC the nature of the *top* quark is expected to be tested to an excellent precision. In particular, for each fb^{-1} of integrated luminosity about 15000 events are expected in the $t\bar{t} \to (l\nu b)(jjb)$ decay channels and about 2500 in the $t\bar{t} \to (l\nu b)(l\nu b)$ decay channels.

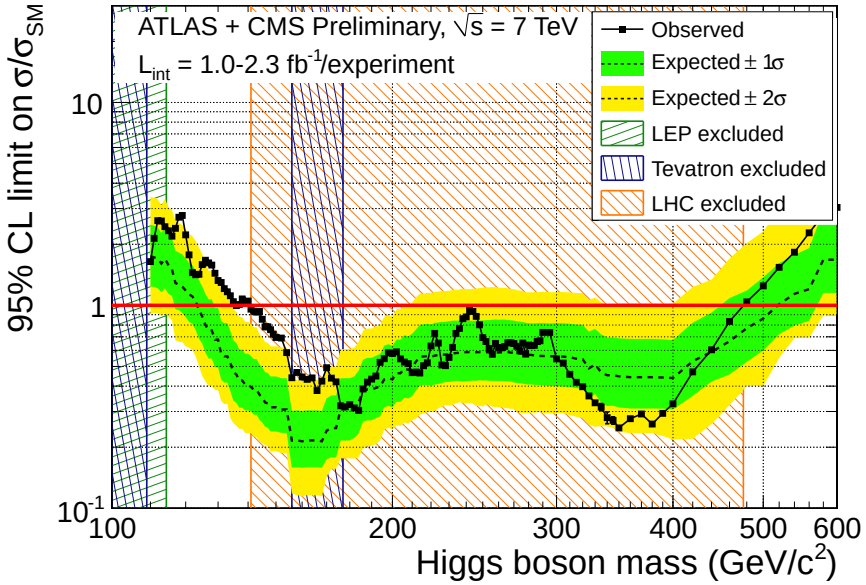


Figure 1.9: ATLAS and CMS combined results for the expected (dashed) and observed (solid) cross section limits at 95% CL for the individual search channels, normalized to the Standard Model Higgs boson cross section, as functions of the Higgs boson mass.

Experimental results

Several works have been published by LHC experiments concerning the *top* properties and the main results of ATLAS and CMS may be found, respectively, in reference [18] and [19]. Among the others, the ATLAS [20] and CMS [21] results on *top* mass measurements are shown, respectively, in figure 1.10 and 1.11: both experiments quote a value in agreement with the expectations and achieved a precision almost comparable to the one achieved by the Tevatron experiments.

1.2.3 B Physics

As for the *top* quark, the b quark cross section is very high at LHC ($\sigma_{b\bar{b}} \sim 300\mu b$) and more than 10^{12} $b\bar{b}$ pairs have been produced so far; hence it will be possible to perform several precision measurements on B mesons physics that will give a better knowledge of many Standard Model parameters, of CP violation in the B meson systems and of possible deviation from the Standard Model predictions in the rare decays.

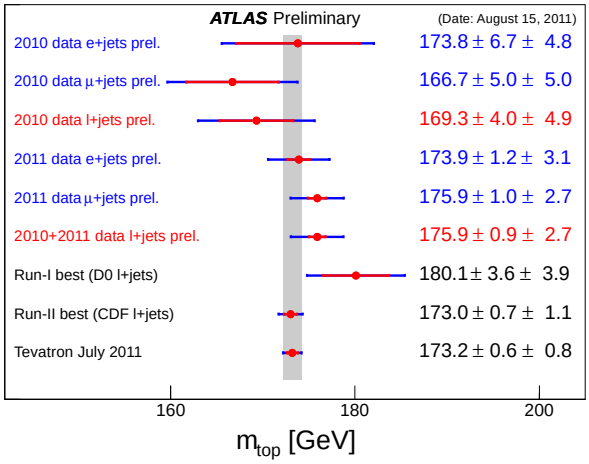


Figure 1.10: ATLAS *top* mass results.

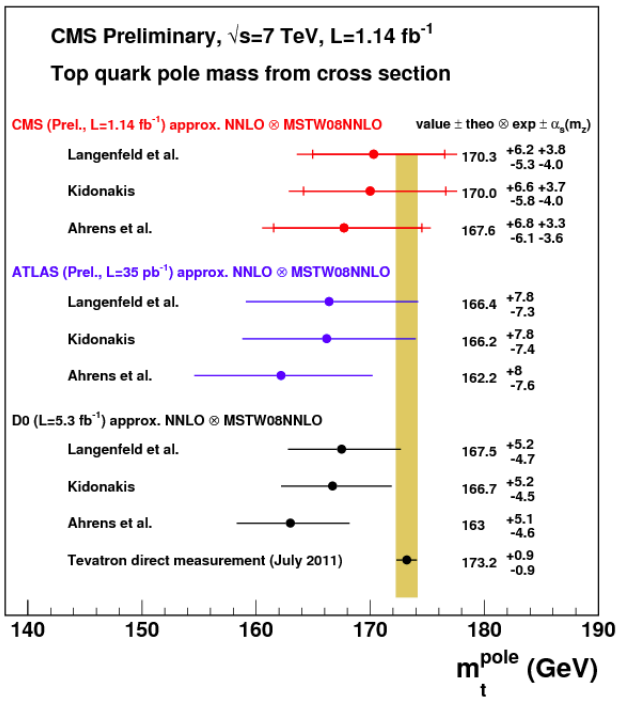


Figure 1.11: CMS *top* mass results.

Although LHCb experiment has been built to perform dedicated studies

in this field and will achieve the best performance in these measurements, ATLAS and CMS experiments are providing measurements on B mesons as well, in order to test kinematics in a different rapidity region [8]. Moreover, as can be seen in figure 1.5, LHCb instantaneous luminosity is lower than the one measured in the other two detectors since, as previously states, this choice allows for an optimization of collision to perform dedicated searches. This means that, especially in the first years of data-taking, when trigger will maintain low selection thresholds until the nominal rate will be reached, the larger amount of data collected will compensate the performance limits of the general-purpose detectors, so the three experiment will provide competitive results.

Experimental results

Besides the precision measurements performed by the LHC experiments, that can be found in [22], [23] and [24] respectively for ATLAS, CMS and LHCb, two main results have been obtained in the B physics searches:

1. the measurement of the ϕ_s parameter [9] in the $B_s \to J/\psi\phi$ decay performed by LHCb [25] with a better precision than Tevatron experiments;

2. the best upper limits on $B_s \to \mu^+\mu^-$ branching ratio.

For what concern the former measurement, within the Standard Model this parameter can be extracted by the following expression

$$\exp^{-i\phi} = \frac{V_{tb}^* V_{td}}{V_{tb} V_{td}^*}, \tag{1.7}$$

where $V_{tb(d)}$ are the CKM matrix elements (see section 2.1.1) that are related to the $t \to b(d)$ transitions. In absence of CP violation $V_{tb(d)}^* = V_{tb(d)}$ and this phase is zero. LHCb results, shown in figure 1.12 are in agreement with the Standard Model predictions although both D0 and CDF obtained a non-zero result but with less accuracy.

Finally, being the main subject of this thesis, the recent experimental results on the $B_s \to \mu^+\mu^-$ searches will be discussed in section 5.3.

[8] LHCb is the only "forward" detector at LHC providing measurements in a rapidity range $y \in [2, 7]$, while ATLAS and CMS measurement are performed in a range $|y| \in [0, 2.5]$.

[9] This parameter is related to the CP violation in the neutral B mesons oscillations; it corresponds to the relative phase between the direct decay amplitude and the decay amplitude via $B_s^0 - \bar{B}_s^0$ mixing.

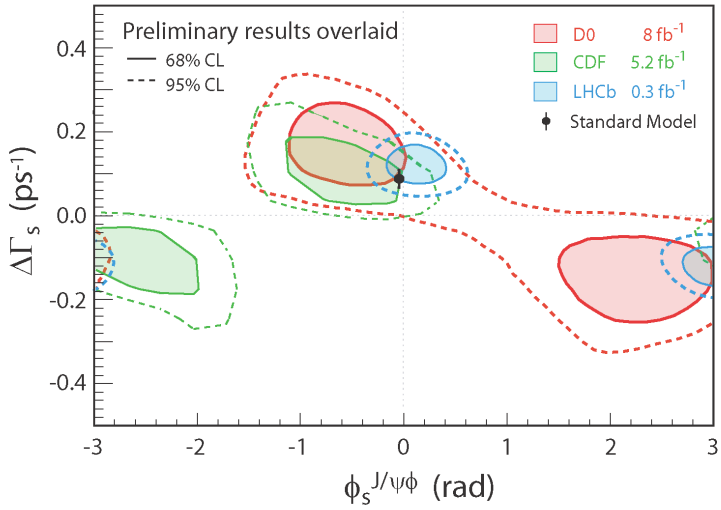


Figure 1.12: LHCb measurement of the decay width difference between the B_s^0 mass eigenstates as a function of the ϕ_s parameter.

1.2.4 Beyond the Standard Model

During the years, high energy experiments experienced an exponential growth, giving the possibility to explore the particle physics spectrum up to the the electroweak energy scale, $M_{EW} \sim 10^3$ GeV.
Nevertheless, between this scale and the the Planck scale M_P there are several orders of magnitude to be explored. Many models have been proposed to fill this gap with other effective theories or to explain the great difference between these scales, the common assumption being the fact that gravity remains unmodified over the ~ 33 orders of magnitude separating the limit of the directly measured gravitational interactions and the Planck length.

The main Beyond Standard Model (BSM) theory being studied at LHC is Supersymmetry (SUSY), which will be briefly delineated in subsection 2.3.1: in the following the latest result on LHC searches will be presented.

LHC constraints on SUSY

SUSY cross sections at LHC energy scale are dominated by the production via strong interactions of gluinos and squarks, that are the supersymmetric partners of gluons and quarks. These particles can give rise to different decay chains ending with the lightest supersymmetric particle, that in many SUSY models is supposed to be massive and non interacting, hence escaping the detection. Together with large missing transverse energy associated to this

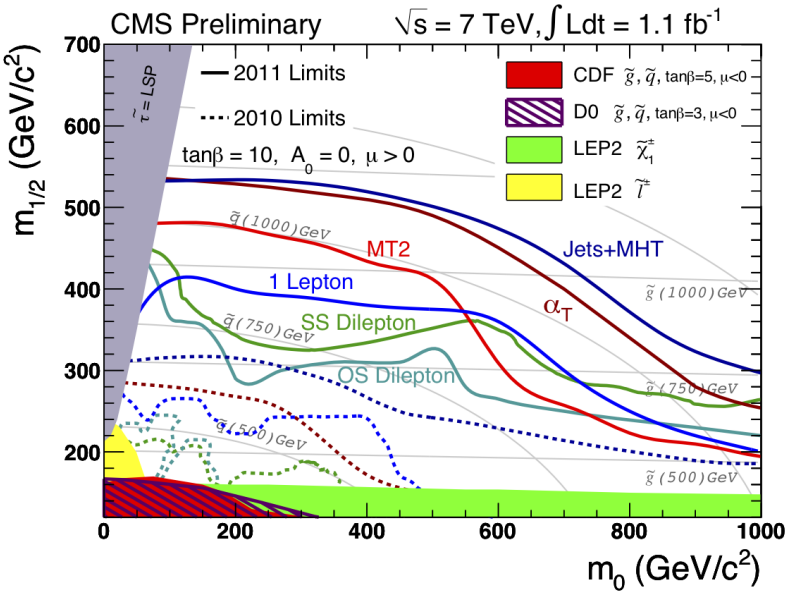


Figure 1.13: Observed limits from several 2011 CMS SUSY searches plotted in the m_0,$m_{1/2}$ plane.

particle, SUSY events contain in the final states hadronic jets and eventually leptons, photons, W and Z bosons; these are distinctive signatures with respect to Standard Model expectations, thus can either prove the SUSY existence or exclude it at LHC energy scale.

A rough estimation of the SUSY mass scale and hence of the accessible phase space in the decays will be given by the peak position of the "effective mass" distribution, defined as the scalar sum of the missing transverse energy and the transverse momentum of the four hardest jets in the events.

If the inclusive searches will be successful more difficult will be to distinguish among the various SUSY models, which are generally characterized by too many free parameters. Since is not possible to investigate all possible cases, a typical approach is to reduce as much as possible the number of parameters and investigate the various scenarios obtained by varying them; generally the various models are expressed by using between four and six parameters related to the masses, to the couplings and to the other relevant variables that may take part in the SUSY decay chains.

Up to now no evidence of SUSY has been found at LHC; the parameter space of one of the main SUSY sub-models, the Constrained Minimal Supersymmetric Standard Model (CMSSM) (cfr. section 2.3.1) has been scanned by looking at the different final state scenarios by both ATLAS [29]

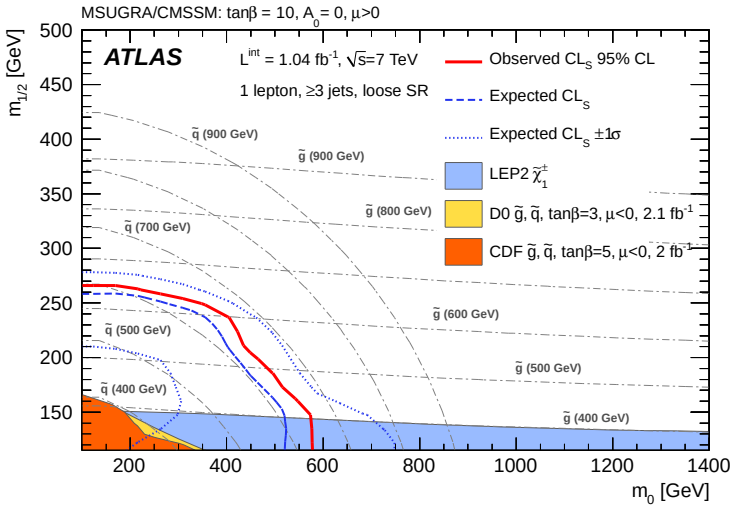


Figure 1.14: ATLAS observed and expected 2011 CL_S exclusion limits at 95% CL for SUSY in the m_0,$m_{1/2}$ plane by using the combined electron and muon channels with missing transverse energy and jets.

and CMS [30] Collaborations. Among the others, the parameters describing this model are m_0 and $m_{1/2}$, which indicate, respectively, the boson masses and the fermion masses at the GUT scale. Examples of exclusions made in this 2D plane by fixing the other parameters can be found in figures 1.13 and 1.14.

Chapter 2

The $B \to \mu^+\mu^-$ decays in Standard Model and Beyond

In this chapter the amplitude for the rare $B \to \mu^+\mu^-$ decays in the Standard Model (SM) and the contributions it may receive from New Physics processes will be illustrated. To this extent the *Flavor Sector* and the Weak Interactions in the SM are introduced; then the effective low energy picture of *Operator Product Expansion* (OPE), useful to give a model independent description of the amplitude, is presented. Finally, an overview of the main Supersymmetric theories that may give deviations of the branching ratios with respect to the SM expectations is given.

2.1 The Standard Model

The whole spectrum of elementary particles is actually illustrated in the Standard Model of Fundamental Interactions (SM) [33, 34, ?]. This is a quantum field theory describing elementary particles, that constitute the excited states of the fields, and their interactions, mediated by gauge bosons. The SM is usually complemented by the Higgs sector [37, 38, 39], although the only particle missing for a full experimental confirm of this picture is the Higgs boson itself. Particles are classified into bosons and fermions depending on the statistics they obey to; fermions (leptons and quarks) are further classified into three generations

$$\text{Leptons: } \begin{pmatrix} \nu_e \\ e \end{pmatrix} \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} ; \text{ Quarks: } \begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}$$

The SM includes three different interactions, corresponding to three conserved charges under local gauge transformations.
Algebraically, it is based on a combination of three symmetry groups:
$SU(3)_C \times SU(2)_L \times U(1)_Y$, representing, respectively, strong, weak and electromagnetic interactions; the associated conserved charges are, respectively, the color, the weak isospin and the hypercharge.

The strong interactions are mediated by eight gluons; in this theory, also known as QCD (Quantum Chromo Dynamics), the coupling constant has an opposite behavior with respect to the QED one: it is small at high energy, explaining the so-called "quark asymptotic freedom" that allows to study them as if they were free, while at small energies increases so much that quarks result "confined" into hadrons.

Electromagnetic interactions are beautifully painted by QED (Quantum Electro Dynamics); the corresponding coupling constant grows with energy and this behavior allows to describe low energy processes by using a perturbative approach achieving an extraordinary precision.
Weak interactions were theorized by Fermi to describe nuclear β decays[40]; this theory was later modified in order to account for parity violation.

$SU(2)_L$ and $U(1)_Y$ groups describe the unified theory of electroweak interactions, mediated by the three vector bosons and by the photon.
There are many differences between weak and electromagnetic interactions: for instance, in the weak case the interaction range is finite and short, while in the electromagnetic case it is infinite [1]. Moreover, neutral currents are allowed in the weak interactions but not in the electromagnetic ones. Nevertheless, in the SM these two theories are unified at an energy scale greater than 100 GeV; at lower energies electroweak interactions arise from the spontaneously broken symmetry $SU(2)_L \times U(1)_Y$ through the Higgs sector.

The electroweak formalism accounts for parity violation by introducing a chiral structure [2] in the currents and in the fields descriptions; thereof, fermion eigenstates can be distinguished in left-handed (L) and right-handed (R)[3]

[1] In a quantum field theory the interaction range is proportional to the inverse of the mass squared of the exchanged boson, so in the weak case it is $\mathcal{O}\left(\frac{1}{100GeV/c^2}\right)$

[2] The invariance under a parity transformation is called chiral symmetry; the violation of parity in weak decays was demonstrated in 1957 in the famous experiment carried out by madame Wu and her collaborators [36].

[3] Fermions fields result eigenstates of the projection operators $\mathrm{P}_{L/R} = \left(\frac{1 \mp \gamma_5}{2}\right)$, where $\gamma_5 = \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}$ and I is the 2x2 identity matrix.

$$L_L = \begin{pmatrix} \nu_{l_L} \\ l_L \end{pmatrix}; \quad L_R = l_R, \nu_{l_R}; \quad Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}; \quad Q_R = u_R, d_R;$$

where l indicates e, μ and τ leptons, u indicates the up-type quarks (u, c, t) and d the down-type ones (d, s, b). Moreover, the V-A (vectorial-axial) structure of electroweak interaction groups the left-handed components in weak isospin $SU(2)_L$ doublets; the corresponding right-handed components are singlets under an $SU(2)_L$ transformation.
Using Gell-Mann and Nishijima relationship[4], the quantum numbers of particles so classified are reported in table 2.1.

Fermions	T	T_3	Y	Q
ν_{l_L}	1/2	+1/2	-1	0
l_L	1/2	−1/2	-1	-1
ν_{l_R}	0	0	0	0
l_R	0	0	-2	-1
u_L	1/2	+1/2	+1/3	+2/3
d_L	1/2	−1/2	+1/3	+2/3
u_R	0	0	+4/3	+2/3
d_R	0	0	−2/3	−1/3

Table 2.1: Quantum number of fermions in Standard Model

The theory so far described is governed by the lagrangian:

$$\mathcal{L} = \mathcal{L}_{(SU(3)_C)} + \mathcal{L}_{(SU(2)_L \times U(1)_Y)} + \mathcal{L}_{(Higgs)} \tag{2.1}$$

where

$$\mathcal{L}_{(Higgs)} = [iD^\mu \phi]^\dagger [iD_\mu \phi] - V(\phi)$$

being ϕ the complex doublet $\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$

and

$$V(\phi) = \mu^2 |\phi|^2 + \lambda |\phi|^4 , \tag{2.2}$$

where $\mu^2 < 0$ and $\lambda > 0$.

[4] $Q_{em} = T_3 + \frac{Y}{2}$, where Q_{em} is the electric charge, T_3 the third component of weak isospin and Y the hypercharge.

Electroweak symmetry breaking occurs when one of the four fields, for instance ϕ_3, is not anymore invariant on the vacuum but acquires a non-zero vacuum expectation value v:

$$\phi_1 = \phi_2 = \phi_4 = 0; \quad \phi_3^2 = -\frac{\mu^2}{2\lambda} \equiv v^2.$$

This leads to the Higgs doublet symmetry breaking:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \to \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}$$

where H is the Higgs boson.
A consequence of the electroweak symmetry breaking is that the physical fields corresponding to the photon and to the three massive vector bosons Z^0, W^+ e W^- appear [41], while the fermions acquire masses through the interaction with the Higgs field. The Lagrangian term corresponding to these interactions for the quarks reads:

$$\mathcal{L}_{Yukawa} = \sum_{j,k} \left[Y^u_{j,k} \bar{u}^j_L u^k_R + Y^d_{j,k} \bar{d}^j_L d^k_R + h.c. \right] \frac{1}{\sqrt{2}} (v + H(x)) \tag{2.3}$$

where $Y^{u,d}_{i,j}$ are the Yukawa matrices; mass terms arise from the terms proportional to v. It is important to notice that these non-diagonal, 3×3 and complex matrices are the only sources of *Flavor* violation in the SM because they mix quark generations. Electroweak currents, in fact, involve mass eigenstates rather than *Flavor* eigenstates: the relation among the two will be deepened in the following section.

2.1.1 The electroweak interactions and the CKM

To determine the quark mass eigenstates it is necessary to diagonalize the mass matrices:

$$m^{u,d}_{i,j} = -\frac{v}{\sqrt{2}} Y^{u,d}_{i,j} \tag{2.4}$$

This means that to connect the electroweak states of quarks with their mass eigenstates, the Yukawa interaction terms induce the unitary transformations:

$$d^{i,mass}_{L,R} \to d^{i,weak}_{L,R} = \sum_j \mathcal{T}^{ij,d}_{L,R}\, d^{j,mass}_{L,R}, \quad u^{i,mass}_{L,R} \to u^{i,weak}_{L,R} = \sum_j \mathcal{T}^{ij,u}_{L,R}\, u^{j,mass}_{L,R} \tag{2.5}$$

where $d = d, s, b$, $u = u, c, t$, and $\mathcal{T}_{L,R}^{u,d}$ are 3×3 matrices performing the diagonalization.

In the weak base the charged interactions of $SU(2)_L$, W boson with quarks contain the product:

$$\bar{u}_L^{weak} d_L^{weak} = \bar{u}_L^{mass} \left(\mathcal{T}_L^u\right)^\dagger \mathcal{T}_L^d d_L^{mass} = u_L^{mass} V d_L^{mass} \tag{2.6}$$

where V is the CKM matrix.
In fact, a generic current charged interaction term reads:

$$\mathcal{L}_{int}^{cc} = \frac{g}{2\sqrt{2}}(W^{+\mu} V_{ij} \bar{u}_L^i \gamma_\mu d_L^j + W^{-\mu} V_{ij}^* \bar{d}_L^j \gamma_\mu u_L^i) \tag{2.7}$$

and the corresponding CP-conjugate term is:

$$\mathcal{L}_{int}^{*cc} = \frac{g}{2\sqrt{2}}(W^{-\mu} V_{ij} \bar{d}_L^j \gamma_\mu u_L^i + W^{+\mu} V_{ij}^* \bar{u}_L^i \gamma_\mu d_L^j); \tag{2.8}$$

note that CP-violation can occur if $V_{ij} \neq V_{ij}^*$.

The CKM is an unitary matrix, schematically represented as follows:

$$V = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} . \tag{2.9}$$

The most useful parametrization of this matrix, expressed in terms of the irreducible phase parameter δ_{13}, is:

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta_{13}} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta_{13}} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta_{13}} & c_{23}c_{13} \end{pmatrix}, \tag{2.10}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$, θ_{ij} represent the Euler angles and i, j are family labels; the main feature of this parametrization is that the family labels are introduced in such a manner that the mixing between two chosen families vanishes if the corresponding mixing angle is set to zero.
It is important to underline that there is a hierarchy between CKM matrix elements[5], i.e.

$$|V_{ub}|^2 \ll |V_{cb}|^2 \ll |V_{us}|^2 \ll 1 \tag{2.11}$$

that can be expressed in terms of Cabibbo angle, being $|V_{us}| = sin\theta_c = 0.22$.

[5]This hierarchy comes out from two experimental observation: the B mesons lifetime, an order of magnitude greater than expected, and the predominance of $b \to c$ decays with respect to $b \to u$ ones.

A key prediction of the Glashow-Weinberg-Salam model was the existence of neutral electroweak interactions. Since these interactions involve particle-antiparticle pairs, CKM matrix does not appear in the currents and *Flavor* results a conserved quantum number. As a consequence, *Flavor Changing Neutral Currents* (FCNC) are highly suppressed; charged current processes are, in fact, mediated by $W^{\pm}$ bosons and their amplitudes result small because involve

- the weak gauge coupling constant g, that is related to the Fermi constant G_F and to the W boson mass M_W through the relation:

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}, \tag{2.12}$$

- the non diagonal CKM matrix elements.

2.1.2 *Flavor Changing Neutral Currents*

Weak interactions processes at the tree level obey certain selection rules:

- *Flavor* changing processes (ΔF=1) involve only charged interactions,
- strangeness quantum number can change only by one unit (ΔS=1).

This means that since FCNC may occur only in higher order processes involving loops, they result particularly useful for both testing the quantum structure of the theory and in the search of the New Physics. The Glashow-Iliopoulos-Maiani (GIM) mechanism, introduced in 1970 [42], accounted for the suppression of these transitions by extending the Cabibbo [43] model and predicting the existence of the charm quark.
The one-loop processes which mediate FCNC can be classified as electromagnetic, weak, or gluonic penguin diagrams and box diagrams; these diagrams can be described by a set of basic triple and quartic effective vertices [44] involving either quarks (penguins) or both quarks and leptons (box), and can be expressed as function of $x_i = m_i^2/m_W^2$, where i indicates the particle involved, and of the CKM parameters. An example of these one-loop diagram is shown in figure 2.1, and the corresponding amplitude can be expressed as follows:

$$\mathcal{A}(b \to s) \propto V_{ub}V_{us}^* x_u + V_{cb}V_{cs}^* x_c + V_{tb}V_{ts}^* x_t =$$
$$V_{tb}V_{ts}^*(x_t - x_c) + V_{ub}V_{us}^*(x_u - x_c) \propto \sin^2\theta_c (x_t - x_c) + \sin^4\theta_c (x_u - x_c). \tag{2.13}$$

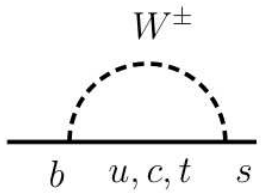


Figure 2.1: One-loop diagram for the b → s transition.

Note that this amplitude would vanish either if the CKM matrix would have a trivial structure, not allowing changing between quark generations, or if the up-quark masses would be identical; moreover this expression in power of $\sin\theta_c$ reflects the small value of the amplitude and the suppression of the process.

The rare decays of B mesons are dominated by Z-boson penguin and box diagrams involving top quark exchange; in particular, for the $B_s \to \mu^+\mu^-$ the main diagrams contributing to the amplitude are reported in fig 2.2. Being this a purely leptonic decay, the theoretical description results quite simple and attractive; in fact, it gives access to the strong interactions binding the quarks in the initial state meson, but no strong interactions are present in the final state, characterized by a rather clear experimental signature. Let's now

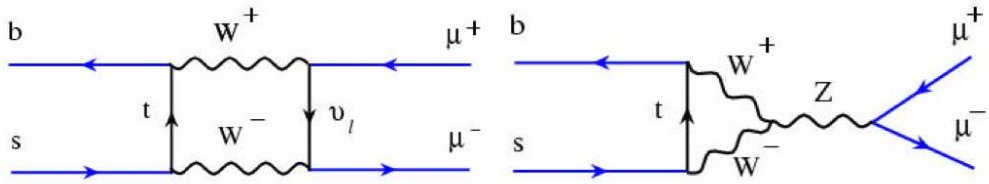


Figure 2.2: Box (a) and penguin (b) diagrams for the $B_s \to \mu^+\mu^-$ decay.

discuss a model-independent description of the mesons decay amplitudes.

2.2 The effective Hamiltonian approach

The strong interactions binding the quarks into hadrons, are characterized by a typical energy scale of ∼ GeV; this relatively low energy allows the description of the leptonic decays of hadrons, which are directly related to the weak interaction between their quark constituents, through the Hamiltonian

[45]:

$$\mathcal{H} = \frac{G_F}{\sqrt{2}} V_{ij} V_{kl}^* \left[\bar{q}_l \Gamma^\mu q_k\right] \left[\bar{q}_j \Gamma_\mu q_i\right] \ , \tag{2.14}$$

where G_F is the Fermi coupling constant, V_{ij} the CKM matrix elements, $\Gamma^\mu = \gamma^\mu(\frac{1-\gamma_5}{2})$ is the V-A current structure, q the relevant quarks involved. The tree-level diagram describing this kind of decays is reported in fig. 2.3.

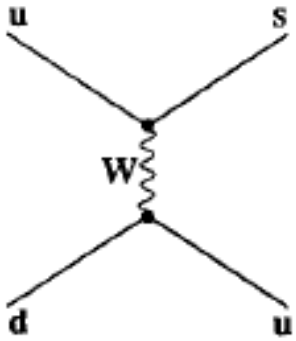


Figure 2.3: Example of tree diagram for the weak interaction between quarks.

Strong interactions affect this W-boson exchange in different ways, depending on the energy of the exchanged gluons with respect to a scale Λ_{QCD}: short range interactions, characterized by an energy scale greater than Λ_{QCD}, can be described by perturbative QCD (pQCD); on the other hand, long range interactions where the energy scale is less or equal to Λ_{QCD}, cannot be treated with a perturbative approach. In order to deal with these complicate effects one starts with an *Operator Product Expansion* (OPE) and then performs a *Renormalization Group Equation*x (RGE) analysis. Using the operator product expansions, the amplitude for the generic $B \to f$ decay can be written as:

$$\mathcal{A}(B \to f) = -\frac{G_F}{\sqrt{2}} V_{ij} V_{kl}^* \sum_\alpha C_\alpha \langle f|O_\alpha|B\rangle \left[\mathbf{1} + \mathcal{O}\left(\frac{m_b^2}{M_W^2}\right)\right]. \tag{2.15}$$

where C_α are the Wilson coefficient [46], O_α are "operator product expansions" for product of four local fields (q_i, $i = 1, ..4$) near the same point [6]. The essential advantage in using the OPE is that it separates the non-perturbative

[6] $O_\alpha = (\bar{q}_1 \Gamma^a q_2)\,(\bar{q}_3 \Gamma_a q_4)$ where $\Gamma_a = (I, \gamma_5, \gamma^\mu, \gamma^\mu\gamma_5, \sigma^{\mu\nu})$ and a indicates the current structure (scalar, vectorial,etc...).

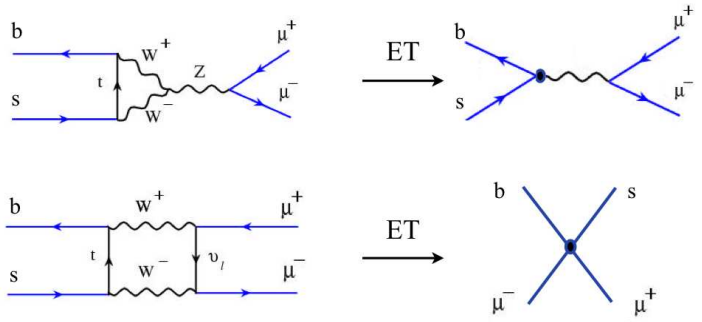


Figure 2.4: Penguin and box diagrams for the $b \to s\mu^+\ \mu^-$ transition from Standard Model (left) to Effective Theory(right) framework.

long-distance contributions, contained in the operator matrix elements, from the short-distance physics, described by the Wilson coefficients; the set of operators needed to describe a specific process depends on his *Flavor* structure. As an example, figure 2.4 shows the relevant diagrams for the $b \to s\mu^+\mu^-$ transition in the Standard Model and in the Effective Theory framework.

The Wilson coefficient can be seen as model independent coupling constant for the effective Hamiltonian:

$$\mathcal{H}_{eff} = \frac{G_F}{\sqrt{2}} V_{ij} V_{kl}^* \sum_{\alpha} C_\alpha(\mu) O_\alpha(\mu) + h.c. \tag{2.16}$$

where μ is the renormalization energy scale[7].
In general the hadronic matrix element ($\langle f|O_\alpha|B\rangle$) entering in the amplitude 2.15 constitutes the major source of uncertainties and is evaluated with non perturbative methods at the μ scale. An amplitude calculated from this hamiltonian, defined at an energy scale of order of the B mass, in principle reproduces the corresponding Standard Model result up to corrections of $\mathcal{O}\left(\frac{m_b^2}{M_W^2}\right)$; more in general, the effective hamiltonian description holds in SM as well as in some SM extensions and New Physics contributions to the amplitudes can either be included in the Wilson coefficients or in the contribution of new operators.

[7]The dependence from the scale μ of C_α assures that the theory is independent from the renormalization scale at which O_α are expressed. If QCD were neglected, they would have the trivial form $C_1 = 1$ and $C_2 = 0$.

2.2.1 The $B_{d,s} \to \mu^+\mu^-$ decays

As already stated in the previous section, the $B_q \to \mu^+\mu^-$ decays are purely leptonic neutral decays proceeding through electroweak penguin diagrams with Z^0 exchange as well as W box diagrams. Since the final lepton-lepton system do not interact strongly, it's possible to factorize the matrix element into an hadronic (J^α_{had}) and a leptonic (J^α_{lept}) current, so that the amplitude will be:

$$\mathcal{A}(B_q \to \mu^+\mu^-) \sim \langle 0|J^\alpha_{had}|B_q\,(p)\rangle \langle \mu^+\left(p^+\right)\mu^-\left(p^-\right)|J^\alpha_{lept}|0\rangle, \tag{2.17}$$

where p, p^+ and p^- indicate, respectively, the four-momentum of the B meson and of the final state muons.
The effective hamiltonian in this case reads:

$$\mathcal{H}_{eff} = -4\frac{G_F}{\sqrt{2}} V_{tb}V^*_{tq} \sum_\alpha C_\alpha(\mu) O_\alpha(\mu), \qquad q = s, d; \tag{2.18}$$

where the relevant operators O_α entering in the sum are:

$$O^q_P = (\bar{q}\gamma_5 b)(\bar{\mu}\gamma_5\mu), \quad O^q_S = (\bar{q}\gamma_5 b)(\bar{\mu}\mu), \quad O^q_A = (\bar{q}\gamma^\alpha\gamma_5 b)(\bar{\mu}\gamma_\alpha\gamma_5\mu), \tag{2.19}$$

Note that the three operators contributing reflect the factorization prescription, while the labels P, S and A denote, respectively, the pseudoscalar, scalar and axial character of the currents.

The branching ratio for the decay $B_q \to \mu^+\mu^-$ ($q = s, d$) is given by [47]:

$$Br(B_q \to \mu^+\mu^-) = \frac{G_F^2\alpha^2 m_{B_q}\tau\,(B_q)}{16\pi^3}\left|V_{tb}V^*_{tq}\right|^2$$
$$\times\sqrt{1-4\frac{m_\mu^2}{m_{B_q}^2}}\left[\left|\mathcal{F}^q_P + 2m_\mu\mathcal{F}^q_A\right|^2 + \left(1-4\frac{m_\mu^2}{m_{B_q}^2}\right)\left|\mathcal{F}^q_S\right|^2\right] \tag{2.20}$$

where $\tau\,(B_q)$ is the B_q meson lifetime, $\mathcal{F}_{A,P,S}$ are Lorentz-invariant functions depending on the Wilson coefficients and on the decay constant F_{B_q}, defined by:

$$\langle 0|\,(\bar{q}b)_{V-A,\mu}\,|B_q\,(p)\rangle = iF_{B_q}p_\mu, \quad (\bar{q}b)_{V-A} = \bar{q}\gamma_\mu\,(1-\gamma_5)\,b \tag{2.21}$$

The functions $\mathcal{F}_P$ and $\mathcal{F}_S$ can be expressed in terms of the Wilson coefficients of scalar and pseudoscalar operators, which within SM are suppressed by a factor $(m_q m_l)/m_W^2$. Thus in equation 2.20 at the leading order $\mathcal{F}_A$ term

dominates. In the case of the B_s, the branching ratio can be shortened as follows [44]:

$$Br(B_s \to \mu^+\mu^-) = 3.4{\cdot}10^{-9} \left[\frac{\tau\,(B_s)}{1.6ps}\right] \left[\frac{F_{B_s}}{210MeV}\right] \left[\frac{|V_{ts}|^2}{0.040MeV}\right] \left[\frac{\bar{m}_t\,(m_t)}{170GeV}\right]^{3.12} \tag{2.22}$$

and a similar expression holds for $B_d \to \mu^+\mu^-$ decay; the foreseen SM branching ratios for the two decays are:

$$Br\left(B_s \to \mu^+\mu^-\right) = (0.32 \pm 0.02) \times 10^{-8}, \tag{2.23}$$

$$Br\left(B_d \to \mu^+\mu^-\right) = (0.010 \pm 0.001) \times 10^{-8}. \tag{2.24}$$

The CDF collaboration has recently announced a possible first signal for the $B_{d,s} \to \mu^+\mu^-$ [48] exceeding the SM predictions, quoting a $Br\,(B_s \to \mu^+\mu^-) = \left(1.8^{+1.1}_{-0.9}\right) * 10^{-8}$, although with a low significance, while latest LHC experimental results from LHCb [49] and CMS [50] do not confirm this excess, being the combined upper limit obtained by the two collaborations of $Br\,(B_s \to \mu^+\mu^-) < 1.1\times\ 10^{-8}$ at 95% confidence level [51].

2.3 Beyond Standard Model

The SM provides a very predictive description of the elementary particles and the Higgs boson results the only missing part of this picture. Nonetheless, there are several open issues from both theoretical and phenomenological points of view that need to be addressed:

- there are at least 25 free parameters to be determined with experiments: 3 gauge couplings, 12 masses (quark and leptons), 6 mixing angles, 2 CP-violating weak phases and 2 parameters to characterize the Higgs sector;
- there are 16 order of magnitude in the difference between the electroweak scale $M_{EW} \sim 100$ GeV/c^2, which defines the scale of electroweak spontaneous symmetry breaking, and the Planck scale, at which Gravity becomes strong;
- there is no account for Baryogenesis since baryon number B is conserved;
- it does not incorporate the physics of dark energy nor of the full theory of gravitation as described by general relativity.

Different New Physics models have been built to address all or part of these issues, but most of them cannot be proved experimentally with the current energy reach of the accelerator machines. In general, even if the energy scale of the process are too low to reveal new particles beyond SM, they can enter in loop diagrams as virtual states; this render rare decays of mesons particularly attractive since, as we saw in the previous section, the amplitude can be enhanced.

In particular, SM prediction for the $B_{d,s} \to \mu^+\mu^-$ branching ratio can be shortened as follows

$$Br\left(B_q \to \mu^+\mu^-\right)_{SM} \propto \left(\frac{1}{16\pi^2}\right)^2 |V_{tb}^* V_{tq}|^2 \left(\frac{m_B m_\mu}{m_W^2}\right)^2 \tag{2.25}$$

In this expression there are two source of suppressions:

1. the *Flavor* suppression term $|V_{tb}^* V_{tq}|^2$;
2. the Yukawa suppression term $\left(\frac{m_B m_\mu}{m_W^2}\right)^2$.

In some New Physics models these contributions may be enhanced; an enhancement of the first term may hint contributions to *Flavor* conversion, either indicating new sources of *Flavor* violation or a more complicated combination of the CKM matrix elements; an enhancement of the second can instead reveal a more complicated Higgs sector. Unfortunately, if an enhancement will be seen it's not possible to disentangle from which of the two sources it may come. New Physics may also contribute to the amplitude with a totally different functional dependence with respect to SM; this confirms the fact that $B_{d,s} \to \mu^+\mu^-$ decays constitute golden channels to detect the presence of New Physics, but not to distinguish between the various models.

2.3.1 $B_{d,s} \to \mu^+\mu^-$ and Supersymmetry

One of the most explored theory Beyond Standard Model is Supersymmetry (SUSY) [52] which includes the SM as an effective theory. The main issue addressed by this model is the so-called "naturalness" problem of the Higgs mass, which in his fine tuning receive corrections of the order of the Plank mass (10^{18} GeV/c^2), even if it is expected to be of the order of 100 GeV/c^2. This issues is solved by introducing in the fine tuning "counter-terms" contributions, due to the interaction of the Higgs fields with new particles, "superpartners" of SM particles for having the same quantum numbers and masses but obeying to different spin-statistics. In particular, to each SM fermion this new symmetry associates a complex scalar, while to the gauge bosons it

associates a fermion. Obviously since SUSY particles have not been observed yet one assumes that SUSY is broken to an energy scale of the order of 1 TeV or greater.

In some SUSY models FCNC processes can proceed at tree level through the exchange of a supersymmetric scalar particle; these models can be classified as *'non-Minimal Flavor Violation'*, as opposite to the*Minimal Flavor Violating* (MFV) ones. The latter category groups models based on the assumption that *Flavor* and CP violation in the quark sector are dominated even beyond SM by the Yukawa couplings, while in the non-MFV picture this assumption does not hold and there may be new *Flavor* parameters and complex phases.

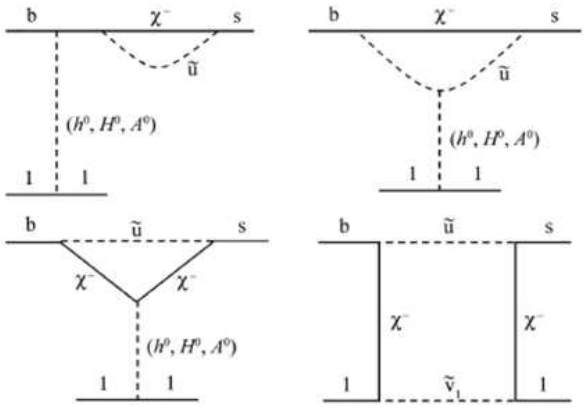


Figure 2.5: MSSM diagrams for the $B^0 \to l^+ l^-$ decays.

One of the possible feature that can distinguish between SUSY models is the conservation of R-parity, defined as

$$R = (-1)^{3(B-L)+2s}$$

where B is the barionic quantum number, L the leptonic quantum number and s the spin.
R-parity conservation has some interesting implications:

1. in colliders SUSY particles must be produced in couples with opposite R-parity;
2. there must be a stable Lightest Supersymmetric Particle (LSP);
3. each SUSY particle but the LSP should decay into a odd number of SUSY particles.

The simplest model conserving R-parity obtained by extending SM is the *Minimal Supersymmetric Standard Model* (MSSM): to all SM particles is associated a superpartner, while the Higgs sector is extended to two complex scalar doublets. Once these two doublets acquire a vacuum expectation value the electroweak breaking occurs.
An high value of the parameter $\tan\beta = v_u/v_d$, corresponding to the ratio of the two Higgs doublets vacuum expectation values, plays a key role in detecting FCNC since branching ratios grow with it; in particular, the $B_s \to \mu^+\mu^-$ rate varies as $\tan^6\beta$. The MSSM Feynman diagrams contributing to the $B^0 \to l^+l^-$ are shown in figure 2.5, while a detailed description of these contributions can be found in [53]. Unfortunately, MSSM contains 105 free parameters, resulting not attractive to test experimentally, so usually models obtained from MSSM by reducing the number of free parameters are preferred; the main distinction between these theories in most of the cases is the sector that operates supersymmetry breaking [8].
By using the latest LHC combined limit on $B_s \to \mu^+\mu^-$ decay and the double ratio [9] built by using the four leptonic meson decays

$$DR = \frac{\Gamma(B_s \to \mu\mu)}{\Gamma(B_u \to \tau\nu)} \frac{\Gamma(D \to \mu\nu)}{\Gamma(D_s \to \mu\nu)}$$

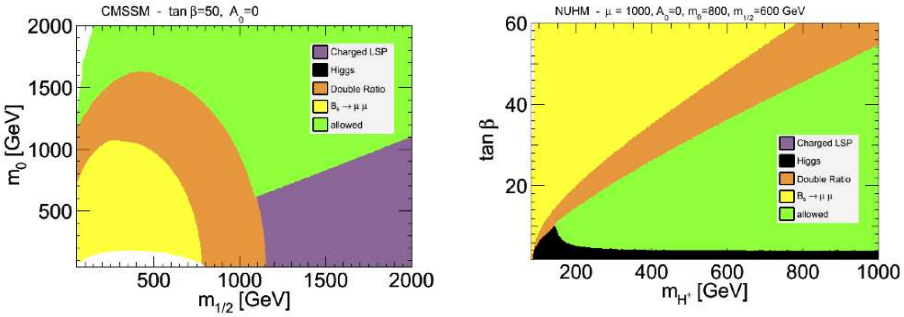
the parameter-space of a set of four SUSY models in the MSSM contest has been investigated [54]:

1. constrained MSSM (CMSSM);
2. non-universal Higgs masses(NUHM);
3. Anomaly Mediated Supersymmetry Breaking (AMSB);
4. Gauge Mediate Supersymmetry Breaking (GMSB).

The CMSSM with supergravity inspired breaking terms is characterized by 5 parameters: m_0, $m_{1/2}$, $\tan\beta$, sign(μ), A_0, where m_0 and $m_{1/2}$ are, respectively, the common masses of bosons and fermions at the GUT (Grand

[8] Symmetry breaking may occur into two different ways, either adding non invariant terms in the lagrangian or through the spontaneous symmetry breaking mechanism. In MSSM Supersymmetry breaking is operated by adding non-invariant "soft" terms that preserve the hierarchy between electroweak and Plank energy scale.

[9] The use of this quantity has an essential advantage since it results independent on the values of the decay constants, that usually constitute the major source of uncertainties in evaluating the decay amplitudes.



(a) CMSSM (b) NUHM

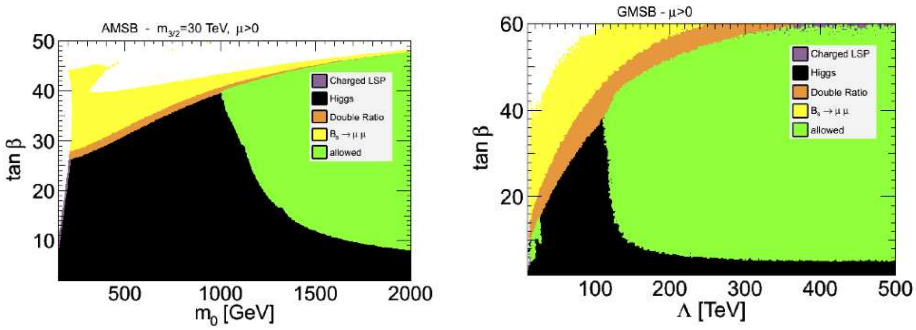
Figure 2.6: Constraints from $B_s \to \mu^+\mu^-$ upper limit and double ratio DR in the CMSSM plane ($m_{1/2}$, m_0) for A_0=0 and $\tan\beta = 50$ (a) and in the NUHM plane (m_{H^+}, $\tan\beta$) for μ =1000, A_0=0, m_0=800 GeV and $m_{1/2}$=600 GeV (b).

Unification Theory) scale, μ is the energy scale at which electroweak symmetry breaking occurs, and A_0 is the trilinear (Higgs-squark-squark or Higgs-slepton-slepton) coupling term. Figure 2.6 (a), shows the constraints obtained in the ($m_{1/2}$, m_0) plane for A_0=0 and $\tan\beta = 50$ by using these two tools.

The second model under investigation, the NUHM, can be considered an extension of the CMSSM since introduces two additional parameters, m_{H_1} and m_{H_2}, corresponding to the mass parameters of the Higgs doublets, that are allowed to be different from m_0. The constraints obtained for this model in the (m_{H^+}, $\tan\beta$) plane for μ =1000, A_0=0, m_0=800 GeV and $m_{1/2}$=600 GeV are shown in figure 2.6(b).

The AMSB is so-named because the soft terms breaking supersymmetry correspond to an "anomalous" violation of a local superconformal invariance, an extension of scale invariance, occurring through a gauge fixing. Graviton superpartner's mass, $m_{3/2}$, is one of the four parameters of the model, the others being m_0, $\tan\beta$ and sign(μ) and having the same meaning as before; the constraints obtained in this case in the plane (m_0,$\tan\beta$) for $m_{3/2}$=30 TeV and $\mu > 0$ are reported in figure 2.7(a).

As for the CMSSM, the GMSB can be described in terms of five parameters: N_5, M_{mess}, Λ, $\tan\beta$ and sign(μ). In this model there are a supersymmetry breaking sector and a "messengers" sector: the interactions between the two operates the SUSY breaking. Messengers usually are grouped into SU(5) mutiplets, and the N_5 parameter indicates the number of copies of the minimal $\mathbf{5} + \bar{\mathbf{5}}$ SU(5) representation. Moreover, M_{mess} is the messengers mass scale while Λ is related to the breaking energy scale; the scan in the



(a) AMSB (b) GMSB

Figure 2.7: Constraints from $B_s \to \mu^+\mu^-$ upper limit and double ratio DR in the AMSB plane (m_0,tan β) for $m_{3/2}$=30 TeV and $\mu > 0$ (a) and in the GMSB plane (Λ, tan β) for $\mu > 0$ (b).

GMSB plane (Λ, tan β) for $\mu > 0$ is reported in figure 2.7(b).

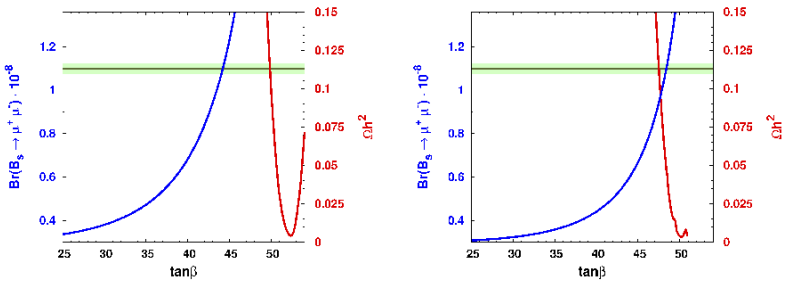


Figure 2.8: The tan β and relic density for A_0=0 (a) and $A_0 > 0$ (b) dependence of $B_s \to \mu^+\mu^-$.

As already stated, and as has been showed in the various MSSM contents, the tan β parameter is strongly correlated to the $B_s \to \mu^+\mu^-$ branching ratio; a very recent study [55] relates it also to the relic density of the Universe Ωh^2[10] and to the direct Higgs constraints obtained at LHC.

Figure 2.8 shows how the $B_s \to \mu^+\mu^-$ branching ratio and the relic density depend on this parameter in the CMSSM scenario for two different assump-

[10]The density of matter in the Universe, ρ, is defined in terms of the ratio $\Omega = \frac{\rho}{\rho_c}$, where ρ_c is the critical density necessary to render the Universe spatially flat. The density of matter is further classified into Light or Dark, the latter being the most abundant one. Different theories relate Dark Matter and SUSY, assuming supersymmetry particle production in the early stage of the Universe.

tions on the A_0 parameter.

From all these results two important messages come out:

- there is still room for New Physics Beyond the Standard Model, since there are still un-excluded regions in the various parameters spaces;
- the $B \to \mu^+\mu^-$ decays constitute golden channels either to reveal or to exclude New Physics so precise measurement of the branching ratio are of primary relevance.

Next chapter will be focused on the ATLAS experimental apparatus and the subsystems that allow to perform precise physics measurements also in the case of the rare decays.

Chapter 3

The ATLAS experiment at LHC

This chapter will explore the detectors, the trigger and data acquisition of the ATLAS experiment; moreover, a brief overview of the performance achieved during the first two years of data-taking will be given.

3.1 The ATLAS detector

ATLAS [2] is a general purpose detector [1] located along the LHC ring whose design has been leaded by the aim to optimize at best, compatibly with the costs, all the sub-detectors in order to achieve good performance in revealing single objects and fulfill the specific aims of the various physics searches, within the limits imposed by the hard LHC environment.

More in detail, the specific requirements that the ATLAS design had to address are the following:

1. perform an efficient revelation of the charged particles tracks, also at high luminosity regimes;

2. provide a complete and detailed event reconstruction;

3. provide an electromagnetic calorimetry able to identify photons and electrons, complemented by the hadronic calorimetry for the jets and missing transverse energy measurement, both providing the full coverage;

4. measure the muon momentum with good precision;

[1] An example of a specific purpose detector, instead, is LHCb; the design of this forward detector has been leaded by the B meson physics event topology B, concentrate close to the line of the beam pipe.

5. select events of interest in the different final state particles energy ranges.

Many of these requirements derive from one of the specific goals of the ATLAS Physics Programme, i.e. the search for the Higgs boson decays into high transverse momentum muons, electrons or photons; a good measurement of missing transverse energy allows to perform search for New Physics, while a trigger system capable to select also low transverse momentum leptons is necessary for B mesons physics studies.

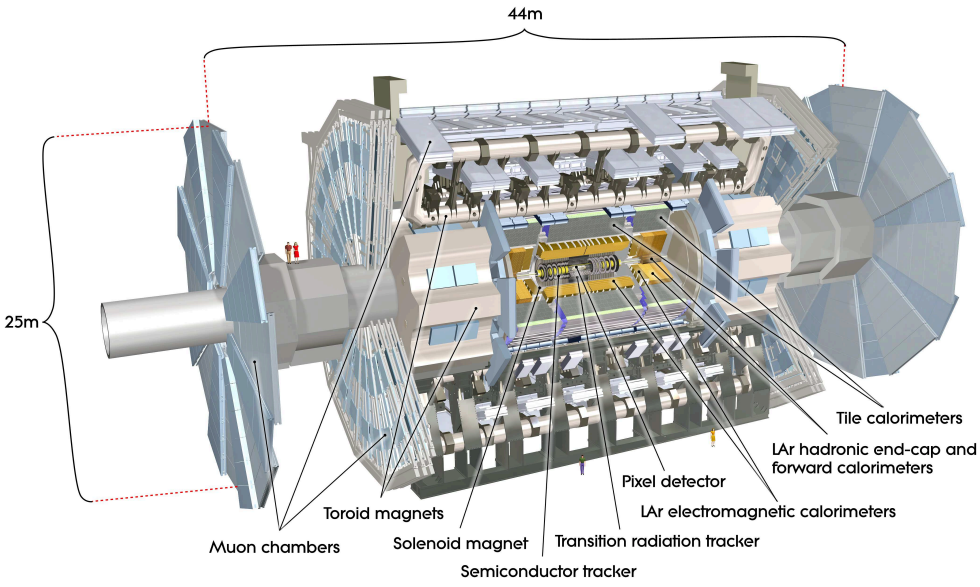


Figure 3.1: Schematic view of the ATLAS detector.

The experimental apparatus layout is schematically represented in figure 3.1: its structure can be divided into three main regions

- the central region, called *barrel*, containing the detectors disposed into concentric layers with respect to the beam axis;
- two *end-caps*, placed on both sides of the barrel.

From the inside out, the detectors composing the barrel can be grouped into

1. the Inner Detector, that contains the tracking system and is surrounded by a solenoidal magnetic field;

Detector system	Required resolution
Tracker	$\sigma_{p_T}/p_T = 0.05\% p_T \oplus 1\%$
EM calorimeter	$\sigma_E/E = 10\%\sqrt{E} \oplus 0.7\%$
Hadronic calorimeter (barrel+endcaps)	$\sigma_E/E = 50\%\sqrt{E} \oplus 3\%$
Hadronic calorimeter (forward)	$\sigma_E/E = 100\%\sqrt{E} \oplus 10\%$
Muon spectrometer	$\sigma_{p_T}/p_T = 10\% p_T$ at 1 TeV

Table 3.1: ATLAS performance goals for the various detector systems.

2. the Electromagnetic and Hadronic Calorimeters, measuring the particle's energies;

3. the Muon Spectrometer, which is immersed in a toroidal magnetic field and provides both tracking and trigger.

The two endcaps, instead, contain the following detectors

1. the Electromagnetic and Hadronic Calorimeters;

2. the Forward Calorimeters, disposed with a small angle with respect to the beam axis;

3. the Muon Spectrometer dipped, as in the barrel, in the toroidal magnetic field.

As we will see in the following subsections, the inner tracking system has a high granularity in order to cope with the high occupancy, energy and radiation due to the collisions; the two calorimeters guarantee high resolution in identifying photons, electrons, jets and missing energy, that are common final signatures to many interesting searches, and excellent transverse and longitudinal sampling; finally, the muon system has a peculiar air-core structure that allows to detect high transverse momentum muons with an extraordinary resolution. Moreover, both Inner Detector and Muon System can provide independent measurement of the transverse momentum, which can be combined together to achieve the best resolution: in fact, ID resolution is better at low p_T while MS one is better at high (> 50 GeV) p_T.

Finally, given the nominal LHC interaction rate of 40 MHz, the detector read-out and the trigger must achieve challenging performance, and the data acquisition, distribution and analysis need to guarantee similar performance. A summary of the performance goals of the various detection systems is reported in table 3.1.

3.1.1 The Coordinate System

The description of the detector components and of the quantities measured by the ATLAS experiment is based on a well defined coordinate system, briefly summarized in the following.

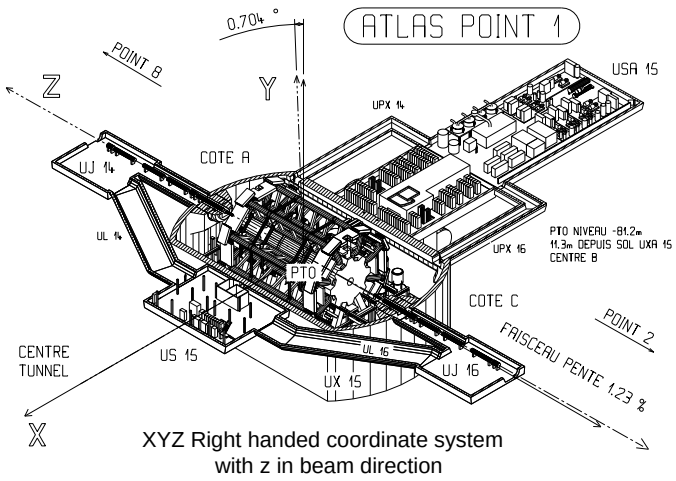


Figure 3.2: Global coordinate system used in ATLAS.

The global [2] coordinate system used within ATLAS is shown in figure 3.2. It is based upon a right-handed, Cartesian coordinate system, where the z axis is oriented counter-clockwise along the beam axis, the positive x-axis points towards the center of the LHC ring and the y-axis points upwards. The origin of the coordinate system is defined as the nominal Interaction Point (IP).

The IP divides the detector into two parts: the A side, corresponding to $z > 0$, and the C side ($z < 0$); the B side corresponds instead to the $z = 0$ plane.

Given the cylindrical geometry of ATLAS, the coordinates are the cylindrical coordinates, i.e the radial distance to the beam axis r, the azimuthal angle ϕ and the polar angle θ, are the most convenient; moreover other useful coordinates are

- the rapidity [3] $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$, where E is the energy and p_z the component of the momentum in the z direction;

[2] The coordinate system is denoted as "global" since all the LHC experiments use the same conventions.

[3] In the hadron colliders this coordinate is preferred with respect to azimuthal angle because the particle's production as a function of the rapidity is roughly constant.

- the pseudorapidity [4] $\eta = -\ln[\tan(\frac{\theta}{2})]$;
- the distance in the $\eta - \phi$ plane, $\Delta R = \sqrt{\Delta^2\eta + \Delta^2\phi}$.

In the following, all the "transverse" quantities in the (x,y) plane will be indicated by the T subscript.

3.1.2 The Tracking System

The ATLAS Inner Detector has been conceived to achieve the best performance in the high occupancy LHC environment; the design criterium has been guided from the following requirements for charged tracks above a given p_T threshold ($\sim 0.5GeV$)

- to have an occupancy $< 1\%$ for each detector element to perform an efficient pattern recognition in both ϕ and z coordinate;
- to supply high tolerance to radiation because of the large flux of particles ($\sim$ 1000 charged particles crossing the sensitive volume of the tracker every 25 ns);
- to minimize the material crossed by particles in order to reduce energy losses and multiple scattering;
- to be able to resolve the production vertices of particles;
- to provide a transverse momentum resolution of $\sigma_{p_T}/p_T = (0.05\% p_T)/\ 1$ GeV $\oplus$ 1 %, while the resolution for high-p_T and very forward particles need to be at least $\sim 30 - 50\%$) ;
- to guarantee high tracking efficiency ($\sim 95\%$) also for minimum ionizing ($p_T \sim 5$ GeV) particles;
- to provide electrons identification having an energy ranging between 1 GeV and 200 GeV over $|\eta| < 2$ by measuring the energy loss (dE/dx).

These goals have been achieved by a careful choice of materials and technologies, compatibly with the budget available.

The Inner Detector surrounds the beam pipe and has a diameter of ~ 2 m, a length of ~ 6 m and provides an angular coverage up to $|\eta| < 2.5$.

[4] This geometric coordinate, that can equivalently be expressed through the relation $\eta = \frac{1}{2} \ln \frac{|p|+p_z}{|p|-p_z}$, approximates the rapidity y in the high energy regimes, when the particle's speed is close to the speed of light c.

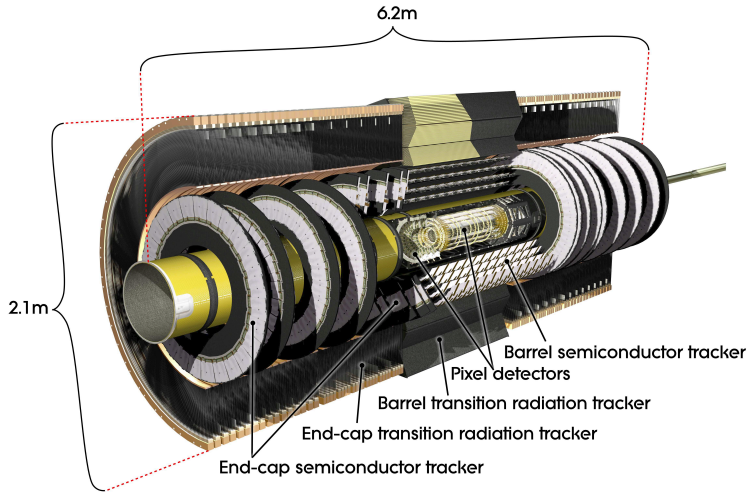


Figure 3.3: The ATLAS Inner detector layout.

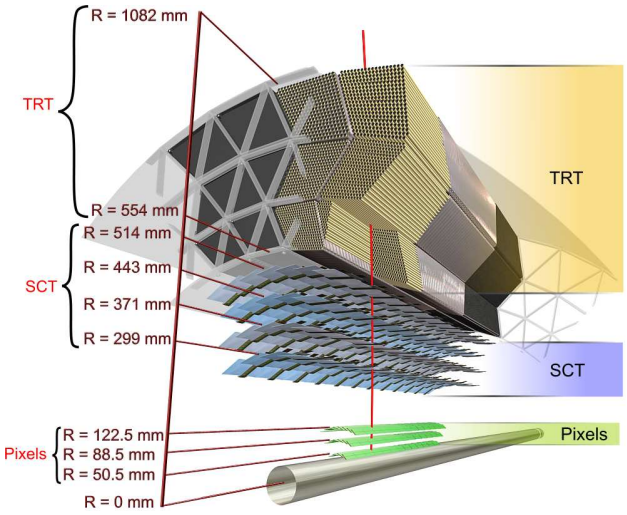


Figure 3.4: Section of the barrel ATLAS Inner detector.

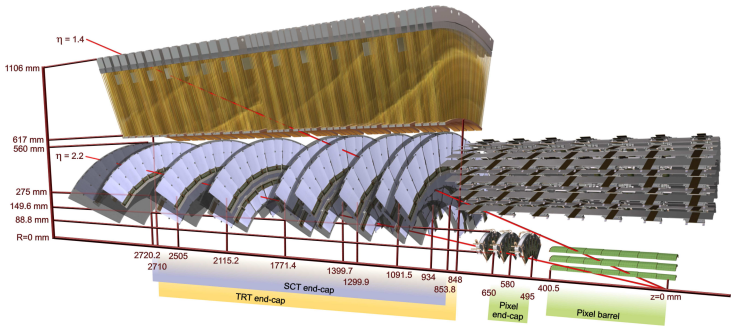


Figure 3.5: Section of the endcap ATLAS Inner detector.

Tracks are bent by a solenoidal magnetic field of 2 T and move along all the detectors through a helicoidal trajectory [5]; the field provided is not uniform since the length of the solenoid is slightly smaller than the tracker.

As is illustrated in figure 3.3, it can be divided into a barrel and two endcaps regions; in the barrel region, from the inside out, there are three technologies, schematically illustrated in figure 3.4

- the silicon Pixel, arranged into three layers disposed, respectively, at a distance r equal to 5.1 cm, 8.9 cm and 12.3 cm with respect to the beam axis,

- the silicon microstrip (SCT), composed of eight layers in the region r$\in [30, 50]$ cm;

- the Transition Radiation Tracker (TRT), disposed in the range r$\in [55, 105]$ cm and composed of 4 mm diameter gas straw tubes that provide almost 35 measured points for each track.

In the endcaps the same detectors are mounted on disks orthogonally with respect to the beam axis (see fig. 3.5): each endcap contains 3 Pixel layers, 9 SCT disks and 4 TRT wheels.

The inner-most detector is made of 1740 rectangular pixels with a minimum $r\phi \times z$ size of $400 \times 50 \mu m^2$, the size being determined by the front-end electronic cells integrated into the sensors by using a *bump bonding*. The detector provides three high resolution tridimensional space-point through the measurement of the charge deposited by each particle crossing it. The

[5] To describe this trajectory five parameters are used: the azimuthal θ and polar ϕ angle, and the impact parameter in the $r-\phi$ and $r-z$ projections, denoted, respectively, as d_0 and z_0, and the curvature radius R, related to the magnetic field B end to the transverse momentum of the particle of charge q through the expression $R = \frac{p_T}{qB}$

Technology	Intrinsic position accuracy in $r-\phi$ (μm)	space-coordinate r and z accuracy (μm)
Pixel	10	115
SCT	17	580
TRT	130	

Table 3.2: Design accuracies for the Inner Detector technologies.

design intrinsic accuracy is strictly related to the size of the pixel elements and to the charge sharing between two adjacent elements.

As the Pixel, the SCT detector provides a measurement of the deposited charge by the particle crossing it. It consist of 4088 modules, each built from two pairs of single-sided silicon micro-strip sensors. Moreover, each side of a module consists of two 6.4 cm long sensors with a strip pitch of 80 μm. In the barrel, stereo strips measure the z coordinate while the ones disposed parallel to the beam axis measure the $r-\phi$ coordinate; in the endcaps, instead, one of the two set is stereo and the other is radially placed.

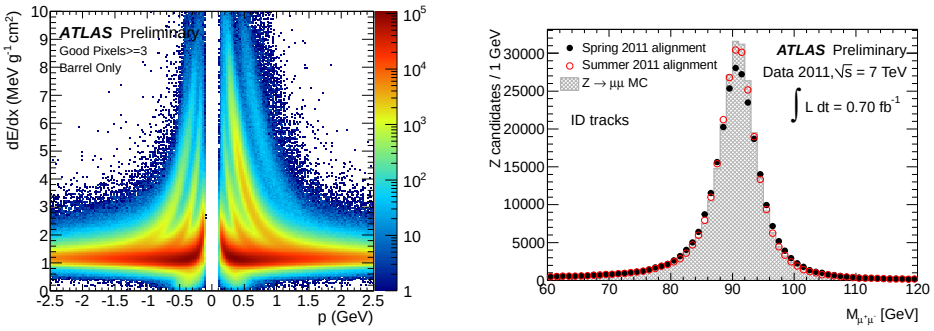
Both Pixel and SCT detectors have been built by using materials having a low coefficient of thermal expansion and are integrated into a robust cooling system [6] in order to cope with the high stability required.

The TRT provide ~ 36 drift time measurement from which the ϕ coordinate of the track is determined. It is made up of ~ 300000 of proportional drift tubes (straws) having a sensing wire in the center. The gas mixture used is composed of 70% Xe, 27% CO_2 and 3% O_2 with 5-10 mbar over-pressure. This mixture allows to use TRT also to perform particles identification by detecting the transition radiation photons emitted and subsequently absorbed in the Xe-based gas mixture. Since the latter yield an electric signal larger than the one produced by the gas ionization due to minimum ionizing charged particles, the distinction between the two signals is obtained by using separate low and high thresholds in the front-end electronics.

Table 3.2 summarize the nominal intrinsic resolutions of the various Inner Detector technologies so far described.

In the following, a brief description of the tracks reconstruction process and of the vertex finding is given; finally, a short overview of the performance achieved with the 2011 data is shown.

[6] Pixel and Microstrip Silicon detectors are kept both at $T \approx -10$ °Cto limit the impact of radiation.



(a) dE/dx measured by Pixel

(b) Z invariant mass from ID measurements

Figure 3.6: (a) dE/dx calculated from the charge collected in the pixel clusters associated to the track of the particles: each curve corresponds to a different identified particle. (b) Invariant mass distribution of $Z \to \mu\mu$ decays reconstructed using track parameters from the Inner Detector track, 2011 ATLAS data are compared to Monte Carlo having an ideal alignment.

Tracks reconstruction and vertex finding

The track reconstruction process begins with the conversion of hits information coming from the detectors into "space-points"; these points are used to find the track seeds from which a primary track direction is determined. From the inside out, hits on the various detector layers are subsequently added to the track fit, taking into account the multiple scattering due to the materials crossed by particles.
Since an event in ATLAS may have several material interaction vertices, an universal vertex finder has been designed to find all vertices in a given track set. In this iterative process, aiming at "cleaning" the vertices set, first all vertex candidates and their associated tracks are found; then the vertices are fitted, taking into account the uncertainty information of all associated tracks, in order to find their positions. Tracks are finally re-fitted with the constraint of coming from the same interaction vertex.

The primary vertex is the one associated to the hardest scattering in the event, the others being considered as pile-up vertices. The former association is done by maximizing the sum of the transverse momenta of the tracks associated to the same vertex.

Some examples of the performance achieved by the Inner Detector are shown in figure 3.6 and can be found, respectively, in references [56] and in [57]. In particular, left plot shows the energy loss dE/dx obtained by measuring the charged released by particles in the Pixel detector by using ATLAS

data. Right plot, instead, shows the Z boson invariant mass built from muon tracks parameters measured in the ID: ideal alignment performance based on Monte Carlo is compared to the observed performance of data processed with spring 2011 alignment and data processed with updated alignment constants; the latter show a very good improvement. More details on the performance of the whole system measured with cosmic-ray data can be found in reference [58]: after the detector alignment, the impact parameter resolutions for high-momentum tracks were found to be $22.1 \pm 0.9 \mu m$ in the transverse direction and $112 \pm 4\ \mu m$ in the longitudinal directions, while the relative momentum resolution was measured to be $\sigma_p/p = (4.83 \pm 0.16) \times 10^{-4} GeV^{-1} \times p_T$.

3.1.3 The calorimeter system

This system provides electromagnetic and hadronic calorimetry, covering the full azimuthal angle around the beam axis and a pseudorapidity range up to $|\eta| < 4.9$); it has been designed in order to determine the particle's energy and the missing transverse energy aiming at reaching a very good energy and angular precision, and linearity in the response. Charged particles identification is provided complemented by the tracker measurements.

More in detail, the design criteria for this system are:

- to perform energy measurements within the designed resolution (cfr. table 3.1);
- to provide high resolution (~ 40 mrad$/\sqrt{E(\text{GeV})}$) in the measurement of the angular coordinates;
- to provide the largest coverage achievable to detect all energy produced in the events and calculate the missing energy;
- to supply good containment for electromagnetic and hadronic showers, limit punch-through into the muon system.

The first two requirements are strictly related to the desirable precision ($\sim 1\%$) in performing Higgs mass reconstruction in the $\gamma\gamma$ decay channel, while the missing transverse energy is an important event feature in many models of Physics beyond the Standard Model, so a good resolution is necessary.

Most of the different subsystems use liquid argon (LAr) as active detector medium, since it provides good linearity in response. The calorimeters closest to the beam line are embedded into three cryostats working at temperatures below 90 $^\circ K$:

- the barrel cryostat, housing the electromagnetic barrel calorimeter, that uses lead (Pb) as absorber ($|\eta| < 1.475$);

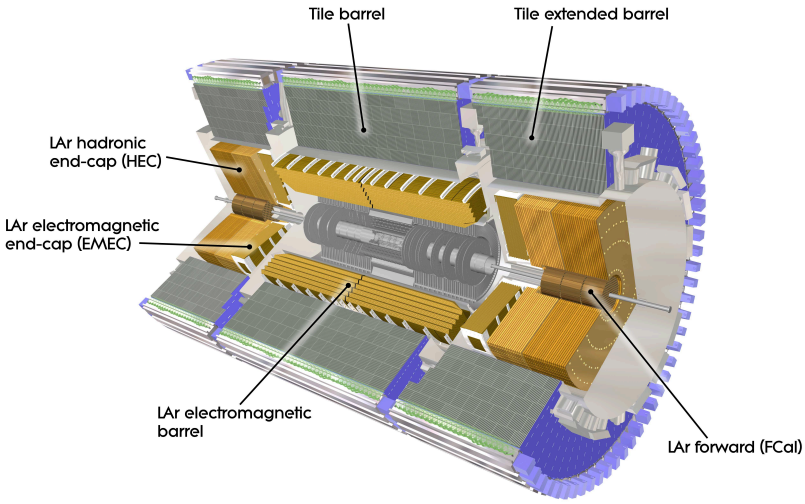


Figure 3.7: The ATLAS calorimeter system.

- two endcaps cryostats, each housing an electromagnetic (EMEC) LAr-Pb calorimeter ($1.375 < |\eta| < 2.5$), a hadronic LAr-copper endcap calorimeter ($2.5 < |\eta| < 3.2$) (HEC) and a forward ($3.1 < |\eta| < 4.9$) copper-tungsten/LAr calorimeter (FCal).

Surrounding the cryostats there is the cylindric hadronic calorimeter (HCAL), made-up of layers of iron-scintillator (used as absorber) and plastic scintillator tiles (as active medium), segmented into a barrel region ($|\eta| <1.0$) and two extended barrel ($1.0< |\eta| <1.7$) regions. Moreover, in the region $|\eta| <1.7$ the electromagnetic calorimeters are complemented by pre-samplers that provide a measurement of the energy lost by particles before reaching them. The layout of this system is shown in figure 3.7.

The electromagnetic calorimeter

The ATLAS EM calorimeter has been built in order to reconstruct the energy of electromagnetic particles like electrons, photons and hadrons.

It presents a peculiar "accordion" shape, that allows a very good and fast sampling with a full ϕ coverage and without azimuthal crack. Honeycomb spacers position the electrodes between the Pb absorber plates, while LAr at $\sim$ 90 °K flows through. The electrodes collect the charge due to the ionization of LAr operated by the shower of particles created in the absorber. The projective geometry allows the calorimeters to have several active layers in

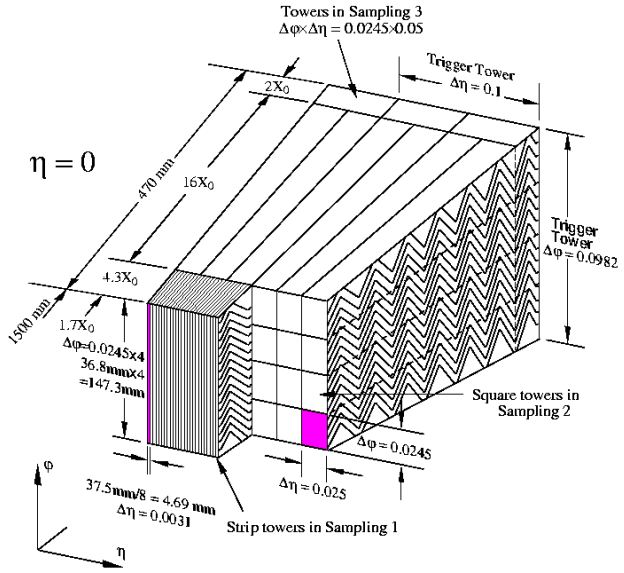


Figure 3.8: Electromagnetic calorimeter cells in the barrel.

depth with different granularity; in the barrel an accurate position measurement is obtained thanks to the fine segmentation of the first layer in η, while the second (middle) layer (*layer 2*) has the highest granularity ($\Delta\eta \times \Delta\phi = 0.025 \times 0.025$) and absorbs the majority of the particle's energy. By grouping the high granular calorimeter cells (see fig. 3.8) one obtains the so-called "trigger towers" that are used in the trigger systems. The total thickness of this calorimeter is greater than 22 radiation length (X_0) in the barrel and above 25 X_0 in the end-caps; the entire system is composed by more than 160000 cells.

The hadronic calorimeter

The hadronic calorimeter uses different technologies in the different pseudorapidity zones. Its thickness has been designed in order to contain the hadronic shower and reduce as much as possible the particles crossing the Muon Spectrometer.

As the electromagnetic, the hadronic calorimeter has different granularity layers but coarser than the electromagnetic one, being the typical cell size of the order of ($\Delta\eta \times \Delta\phi = 0.1 \times 0.1$); moreover, also in this case in the barrel

there are three layers, with the middle one being the thickest.

In the barrel the calorimeter structure is periodic along the z direction: the signals coming from each tile plate serve as input to one photomultiplier while the iron serve as absorber; in the endcaps, instead, the LAr serve as active medium and the copper disks as absorber.

The total thickness is at least 10 absorption lengths (λ) while the total number of readout cells is about 100000.

Forward calorimetry

The forward calorimeters are placed in the same cryostats as the end-cap calorimeters. Being very close to the beam pipe, the FCAL is exposed to high radiation; this lead to a design with LAr gaps smaller than the usual 2 mm gap of the EM barrel calorimeter in order to both avoid ion build-up problem and have faster signals. The FCAL is longitudinally segmented into three parts, where the first, made of LAr-copper, is optimized for EM calorimetry, while in the other two the copper is complemented by tungsten rods with a tubular electrode readout.

Calorimeters performance

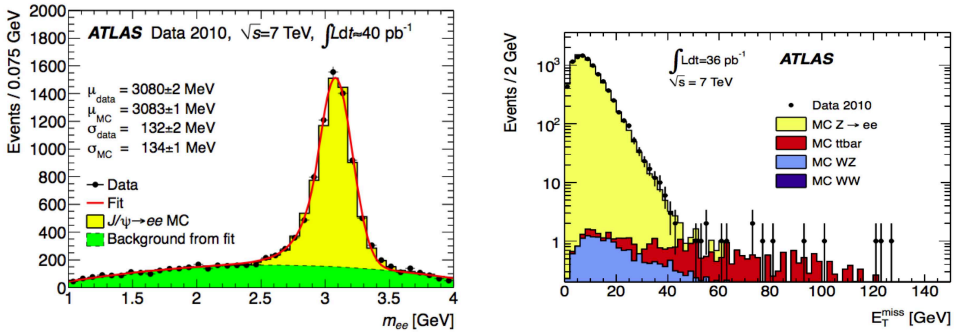
Generally, the energy resolution of a calorimeter is expressed by

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c$$

where a is the sampling term (including statistical fluctuations), b takes into account the noise due to electronics while c it's a constant accounting for systematic uncertainties like the calorimeter non-compensation and all the non-uniformities involved. The ATLAS nominal calorimetry resolution can be found in table 3.1; by using 2010 data [60] the constant term of the EM energy resolution has been measured to be $(1.2 \pm 0.1(\text{stat}) \pm 0.3(\text{syst}))\%$ in the barrel EM calorimeter $\sim$1.8% in the endcaps and to about 3% in the forward regions.

For what concerns the hadronic calorimeter, the fractional energy resolution for isolated pions has been studied [2] as a function of beam energy and impact angle and the value obtained is $\frac{\sigma_E}{E} = \frac{52\%}{\sqrt{E}} \oplus 3\%$, that is in good agreement with the design requirement.

Some example of the very good performance achieved by ATLAS calorimetry can be found in figure 3.9. In particular, left plot [60] shows the $J/\psi \to ee$ invariant mass as obtained by selecting J/ψ candidates in data compared to



(a) $J\psi$ invariant mass (b) E_T^{miss}

Figure 3.9: (a) $J/\psi \to ee$ invariant mass as measured after applying the baseline calibration; data (full circles with statistical error bars) are compared to the sum of the MC signal (light filled histogram) and the background contribution (darker filled histogram). (b) Distribution of E_T^{miss} measured by using $Z \to ee$ candidates in data. The expectation from Monte Carlo simulation is superimposed and normalized to data, after each MC sample is weighted with its corresponding cross-section.

the Monte Carlo samples of signal plus background: the good agreement between data and simulation attests the very good calibration of EM calorimeter. Right plot [61], instead, shows the very good agreement obtained in the comparison of the missing transverse energy (E_T^{miss}) distribution as obtained by selecting $Z \to ee$ candidates and Monte Carlo samples.

3.1.4 The magnets system

The ATLAS experiment has a very peculiar magnets system which names the experiment itself.

It has been designed in order to provide the magnetic field in a large region ($|\eta| < 3$) and allow the inner tracker and the Muon Spectrometer to perform measurement of the charged particles momentum. The ATLAS magnets system is sketched in figure 3.10. This huge system 26 m long with a total diameter of 20 m is composed of

- a central solenoid (CS), providing a 2 T solenoidal field in the tracker;
- three superconducting air-core toroid systems (one in the barrel and two in the endcaps), each made of eight coils, providing a toroidal field up to 4T in the Muon Spectrometer.

The central solenoid is 5.3 m long and has a diameter of 2.4 m; the endcaps toroids placed at both sides of the CS are 5 m long and have an

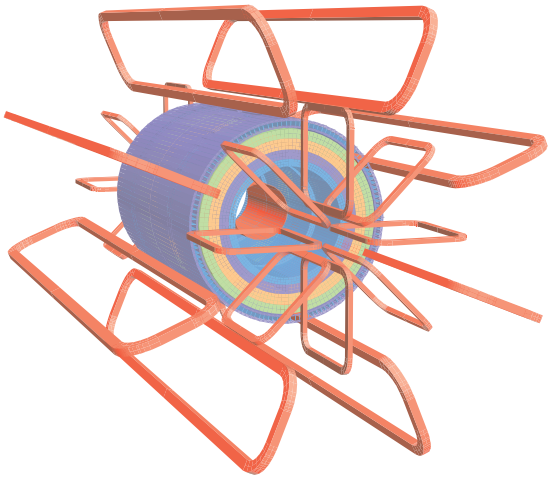


Figure 3.10: The ATLAS magnets system; the toroidal field in the Muon Spectrometer is provided by eight coils in both barrel and endcaps, while embedded in the calorimeters there is the central magnet providing the solenoidal magnetic field to the tracker.

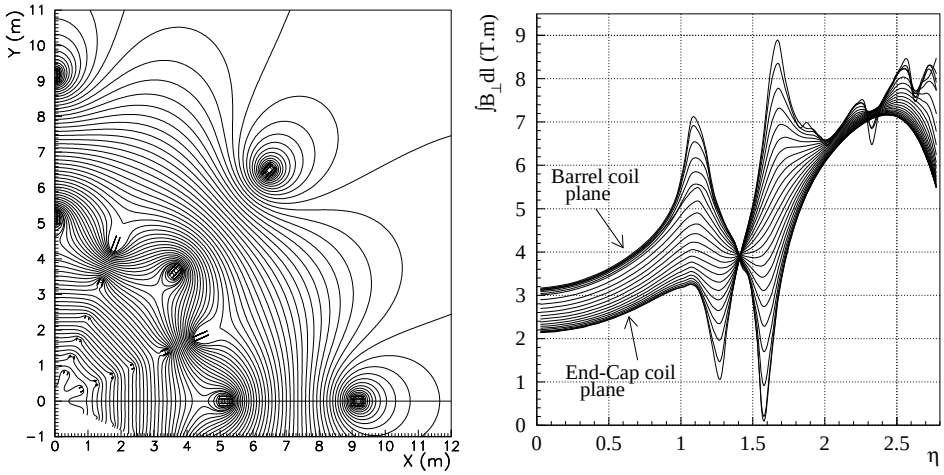


Figure 3.11: (a) Magnetic field lines in the $(x-y)$ plane. (b) bending power as a function of η.

external diameter and an internal bore respectively of 10.7 m and 1.65 m. The endcap coils, assembled radially and symmetrically around the beam axis, are shifted by 22.5°in ϕ with respect to the barrel coils.

Magnets are kept at low temperature by the liquid helium flowing through tubes welded to the coils at a temperature of 4 °K.

The magnetic field B provided is not uniform, as can be seen in the left plot of figure 3.11.

By integrating the azimuthal component of the field along the line that connect the inner bore and the external diameter, one obtains the bending power of the system. The latter is illustrated in the right plot of figure 3.11 as a function of the pseudorapidity; from this plot it's possible to characterize it in different regions

- in the region $0 < |\eta| < 1.3$ the barrel toroids provide a bending power in the range $[2, 6]$ T×m;
- in the region $1.6 < |\eta| < 2.7$ the endcap toroids provide a bending power in the range $[4, 8]$ T×m.

Finally, in the intermediate region ($1.3 < |\eta| < 1.6$) the bending power is lower because of the superposition of the magnets.

The air-core structure of the toroids allows the spectrometer to measure muon momentum with a good resolution since the multiple scattering contribution is minimum, despite the non uniformity of the field.

3.1.5 The Muon System

The Muon Spectrometer (MS) is probably the most innovative ATLAS detector subsystem. It has been designed to detect muons momentum with good resolution on a broad p_T range (∼1 GeV - 1 TeV) and this allows to perform several interesting studies since high energy muons appear as final states in the Higgs boson decays as well as in many New Physics decay chains.

The main features of the MS is that it can both provide a stand-alone trigger system in a wide range of transverse momentum and stand-alone measurements of muon momentum, pseudorapidity and azimuthal angle.

Contrary to the Inner Detector, where charged particles are bent in the $r - \phi$ plane, the toroidal magnetic field bends muons in the $r - z$ plane; this bending [7] allows to measure the momentum with reasonably constant resolution over the pseudorapidity range.

[7] The momentum resolution can be expressed through the expression $\sigma_p/p \sim 1/(BL^2)$, where B is the magnetic field strength and L the lever arm. ATLAS decided to maximize the resolution by using a reasonably magnetic field strength (∼ 0.6 T) but a larger level arm.

The design characteristic of the MS can be summarized as follows

- it allows to reach a transverse momentum resolution of$\sim$ 10% for $p_T \sim$ 1 TeV and of $\sim$ 3% for $p_T \sim$ 100 GeV;
- it has a pseudorapidity coverage up to $\pm$ 2.7;
- it provides a measurement of the η, ϕ coordinates that results into a spatial resolution of $5 - 10$ mm on the off-line reconstructed tracks;
- it guarantees an efficient trigger for both low (< 10 GeV), and high p_T (> 10 GeV) muons;
- it is able to cope with a high background condition [8];
- it provides a good timing and a high resolution in identifying the bunch crossing.

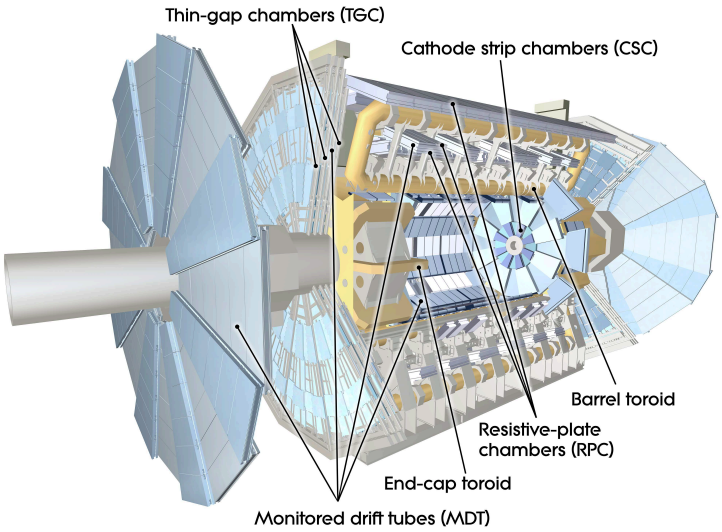


Figure 3.12: Muon Spectrometer view.

As for the other systems, the MS is segmented into a barrel ($|\eta| < 1.05$) and two endcaps ($1.05 < |\eta| < 2.7$). The MS design is sketched in figure 3.12; from this picture it's possible to distinguish four different chamber types:

[8] In the ATLAS cavern the radiation rate is huge ($\sim$ 10 Hz/cm^2 in the barrel region, a few kHz/cm^2 in the forward regions): it is due to neutrons and photons produced by the interaction of the beam protons with the beam pipe or the forward calorimeters. This "cavern background" produces high energy (up to a few MeV) photons that subsequently convert into electrons which cross the spectrometer and are misidentified as muons.

- the Trigger Chambers, composed of Resistive Plate Chambers (RPC) in the barrel and Thin Gap Chambers (TGC) in the endcaps;
- the Tracking Chambers, composed of Monitored Drift Tubes (MDT) complemented by Cathode Strip Chambers (CSC) in the endcaps.

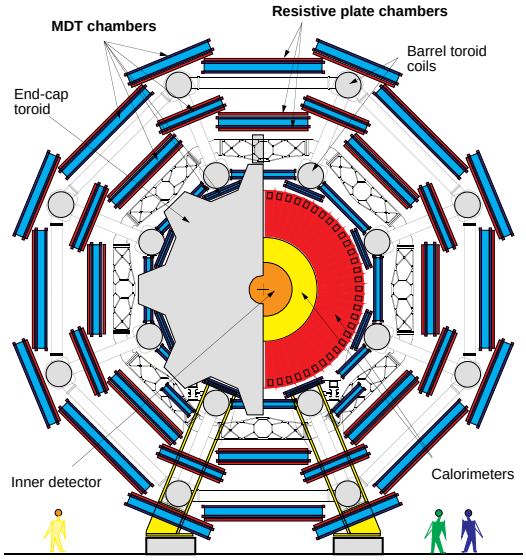


Figure 3.13: Muon spectrometer view in the $x - y$ plane.

In the $r - \phi$ plane the segmentation in both barrel (see figure 3.13) and endcaps follows the octant geometry of the toroids.
Chambers in the barrel are arranged into three cylindrical layers placed at a distance r equal to 5 m, 7.5 m and 10 m with respect to the beam axis; in the endcaps the chambers are arranged in four disks placed at a distance of 7m, 10 m, 14 m and $21 - 23$ m from the interaction point.

This large volume covered with relatively light sensitive materials allows to minimize the multiple scattering effect.

The Muon Trigger in the barrel makes use of three RPC stations each comprising two gas gaps.

In the endcaps ($1.05 < |\eta| < 2.4$) TGCs are arranged in nine layers of gas volumes, grouped into four planes in z, and form two trigger stations: the inner station at $|z| \sim 7$ m consists of one plane of doublet TGC units, while the remaining seven planes, grouped into one plane of triplets and two plane of doublets, form the outer station at $|z| \sim 14$ m.

Both RPC and TGC are fast detector with a time resolution of $\sim 1\ ns$ and provide also spatial coordinate measurements; more details on the MS

chambers will follow.

Monitored Drift Tubes

The MDTs allow for precise tracking all over the MS. In these chambers a spacer frame constitutes the inner support to two multilayers of drift tubes built in aluminum, and since they are oriented orthogonally with respect to the beam axis, they allow the measurement of the $r - \eta$ coordinate. Each multilayer contains three layers of thin walled ($\sim 400\mu m$) tubes in the outer and middle stations and four layers in the inner one. The tube diameter is $\sim$ 3 cm with a length varying from 0.9 to 6.2 m; each tube contains an anode wire 50 μm wide operated at high voltage (3080 V) and a gas mixture of Ar (93%) and CO_2 (7%) that guarantees linearity in the drift time and good ageing properties.

The drift tubes operate in a proportional regime with a maximum drift time of $\sim$ 700 ns and the spatial resolution of each tube is $\sim$ 80 μm. The tubes support is provided by a monitoring structure[9] that avoids the deformation that may occur because of gravity or temperature.

Thanks to this continued monitoring, the accuracy of a track segment in the 6 (8) tube layers is limited only by the single-tube resolution; depending on the number of multilayers, it ranges between 30 and 50 μm.

Cathode Strip Chambers

As MDTs, CSC chambers allow for precision coordinate measurements; thanks to their high rate capability ($\sim$ 1000 Hz/cm^2) and time resolution, they are located in the high pseudorapidity region in combination with MDT chambers.

They are multi-wire proportional chambers employing Ar/CO_2 (80/20) as gas mixture, with an anode pitch of 2.54 mm and a cathode strip read-out with a pitch of $\sim$ 5.31-5.56 mm. The charge interpolation between neighboring strips allows for the precision measurement of the second coordinate ϕ, with a resolution per CSC plane of $\sim$ 60 μm.

The measurement of the η coordinate is obtained either from the signal coming form the wires, that present a maximum drift time of $\sim$ 20 ns, or from the strips placed parallel with respect to the wires on the second cathode plane; in this case the spacing of the readout strips is larger than the one in the bending plane leading ($\sim$ 13-21 mm) to a resolution per plane of about 5 mm.

[9] This structure contains also the RASNIK [65] optical monitoring system that allows the wires position monitoring with a 5 μm accuracy.

Muon momentum reconstruction and performance

The muon momentum in the barrel is obtained by measuring the sagitta [10] s of the muon trajectory inside the magnetic field; the three MDT stations provide three points for this measurement.

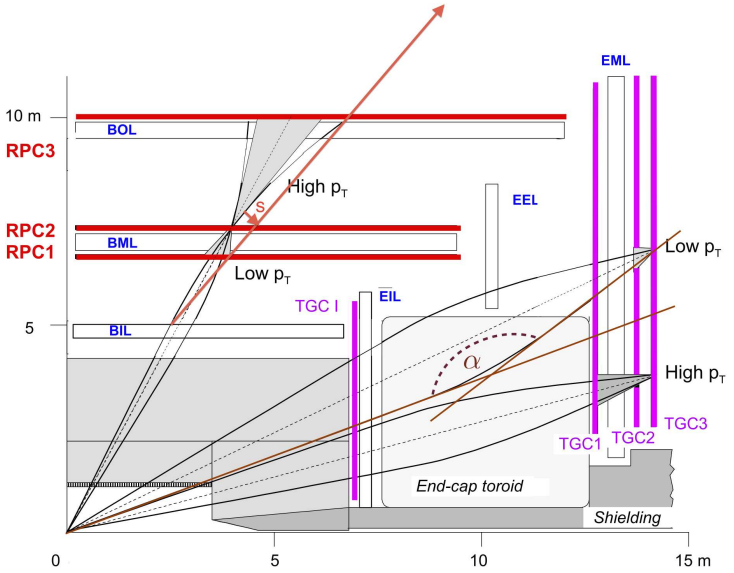
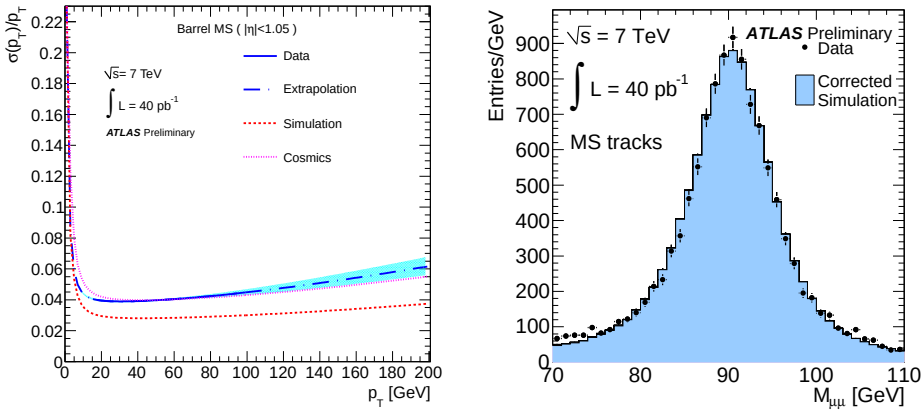


Figure 3.14: Muon momentum reconstruction in the MS

In the endcaps, instead, the muon momentum is measured with a "point-angle" method by using the information coming from the MDT placed on three endcap disks, complemented by the CSC's one in the high rapidity range ($|\eta| > 2$). Because of their geometry, in the endcaps the magnetic field is provided only between the inner and middle stations, so the the momentum is determined by using a point in the inner station and the angle α formed by the intersection of the segment tracks directions in the inner and in the middle-outer stations.

Both reconstruction techniques are sketched in figure 3.14. The accuracy on the muon trajectory measurement reflects in the accuracy of the momentum measurement; moreover, in the latter the effects related to Muon Spectrometer alignment and to the amount of material traversed should be taken into account.

[10] The sagitta of an arc is the maximum perpendicular distance of the arc from the line segment connecting the endpoints.



(a) MS p_T resolution (b) MS p_T

Figure 3.15: (a) Muon Spectrometer resolution in the barrel region: the solid blue line shows results based on 2010 data and is continued as dashed line for the extrapolation to higher p_T. (b) Di-muon invariant mass comparison in the Z boson mass range between collision data (dots) and simulation (full histogram).

More in detail, depending on the muon transverse momentum, the stand-alone resolution [62] of the MS is limited by different contributions:

- for $p_T < 10$ GeV the resolution is dominated ($\sim$ 5-8%) by the fluctuation on the energy loss in the calorimeters;
- for $p_T \in [10, 200]\, GeV$ the multiple scattering affects ($\sim$ 4-6%) the measurement;
- for $p_T > 250$ GeV the larger contribution to the uncertainties is given by the spatial accuracy of the chambers.

Generally, for a given value of η, the momentum resolution can be parametrized as

$$\frac{\sigma(p)}{p} = \frac{p_0^{MS}}{p_T} \oplus p_1^{MS} \oplus p_2^{MS} \cdot p_T \tag{3.1}$$

where the p_i^{MS} coefficients, $i = 0, 1, 2$ are respectively related to the energy loss in the calorimeters material, the multiple scattering and the intrinsic resolution.

Examples of the performance achieved in the muon momentum resolutions evaluated by using 2010 collision data are shown in figure 3.15. Here left plot shows the muon momentum resolution into the barrel region as a function of the transverse momentum compared to the simulation and to the

result obtained by using cosmic rays. This comparison allowed to extract correction factor for Monte Carlo samples, and the result of this correction can be seen in the left plot, where the invariant mass of the Z boson candidates in data results in agreement with the Monte Carlo expectations.

Resistive Plate Chambers

The RPCs present excellent characteristics for a trigger system; in fact, these gas detectors are fast, with a time resolution of ~ 1 ns, and are operated in the "avalanche" regime, that allows to tolerate rates up to $\sim 1\text{kHz/cm}^2$.

Moreover, they provide the ϕ coordinate measurements with a precision of $\sim 10\ mm$; this is fundamental in the muon track reconstruction as MDTs cannot measure the ϕ coordinate.

The chambers are made by two parallel highly resistive ($\rho \simeq 10^{10} \div 10^{11} \Omega$ cm) bakelite plates separated by insulating spacers to form a 2 mm gas gap; the gas mixture is composed of $C_2H_2F_4$/Iso-C_4H_{10}/ SF_6 (94.7/5/0.3) . The SF_6 has been chosen because, even at low concentrations [11], provides a comfortable plateau for the avalanche operation.

High voltage is applied on the resistive plates through the graphite coating of bakelite; the electrical signal is induced by the avalanche on two read out copper strip planes placed on both sides of the gas gap and then amplified and discriminated by fast electronics.

Thin Gap Chambers

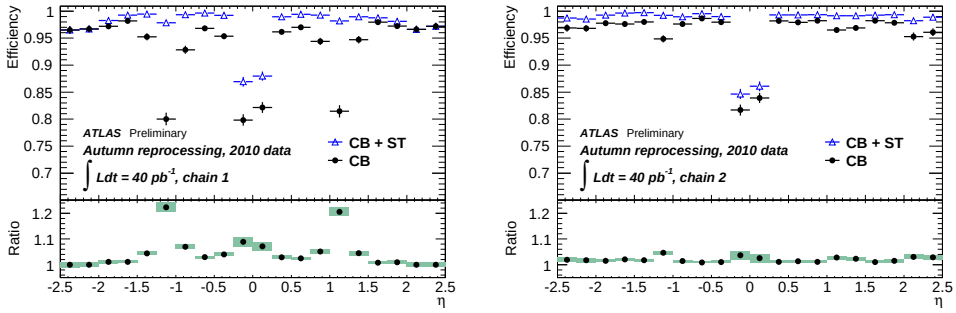
TGC chambers have been chosen as trigger chambers because of their very good rate capability, timing resolution and ageing characteristics.

As CSCs, TGCs are multi-wire chambers with a gas gap of 2.8 mm and a wire pitch of 1.4 mm.
The wires have a diameter of $\sim 50\ \mu$m, with a pitch of 2 mm; they are placed between two graphite cathode planes, each plane provided on the external surface by strips placed either parallel or orthogonally with respect to to the wires direction.

The wires arranged in the ϕ direction provide the trigger signal for the r coordinate, while the strips orthogonal to these wires provide the ϕ coordinate.

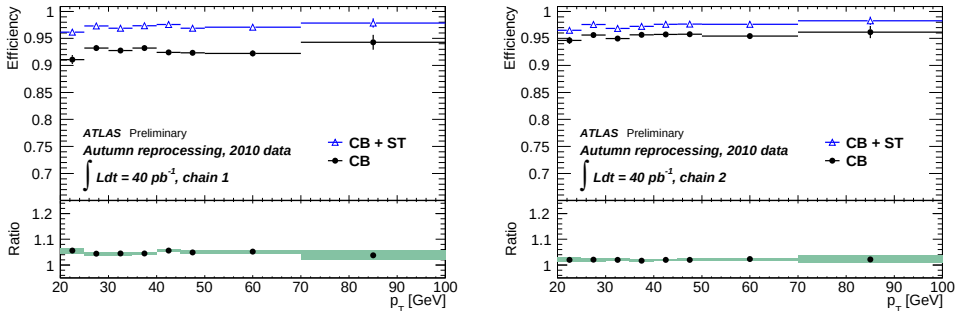
The highly quenching gas mixture used (CO_2/n-pentane, 55/45) allows the operation in saturated avalanche regime with a time resolution of about 4 ns.

[11] The low concentration is dictated also by the aim to avoid the HF formation that degrades the detector.



(a) STACO (b) MUID

Figure 3.16: Muon reconstruction efficiency as a function of the pseudorapidity for the STACO (a) and MUID (b) families: Combined algorithm efficiencies are further improved if complemented by Segment-Tag algorithms.



(a) STACO (b) MUID

Figure 3.17: Muon reconstruction efficiency as a function of the transverse momentum for the STACO (a) and MUID (b) families: Combined algorithm efficiencies are further improved if complemented by Segment-Tag algorithms.

3.1.6 Muon reconstruction algorithms and performance

Two complementary and equally efficient algorithm families perform the offline muon reconstruction within the ATLAS software framework [63]: they are referred, respectively, as STACO (or Chain1) and MUID (or Chain2) muons. Starting from the signals coming from the various detectors, the output of these algorithms results into four different "muon types" depending on the quality of the track associated to the muon. Muon reconstruction process begins with the track finding and fitting: MDT hits on a single muon station are grouped together in straight line called "segment" of tracks, then segments from different stations are combined together to form a Muon Spectrometer track. This track is further extrapolated back to the Interaction

Point, taking into account the multiple scattering and the energy losses in the calorimeters. They are referred to as "Stand-Alone" (SA) muons to spot the fact that they are built using MS only, without any information from the ID. These muons are reconstructed in a large pseudorapidity region ($|\eta| < 2.7$) with some coverage holes due to the absence of muon chambers ($\eta \sim 0$) or to the MS transition region from barrel to endcaps ($|\eta| \sim 1$).

If the SA track parameters match an Inner Detector track, then these two tracks are combined into a "Combined" (CB) muon. With the combination of the two independent information a better resolution and a better rejection power of fake muons coming from in-flight decays is achieved.

The muons with low p_T (< 3 GeV) may not be able to cross the whole spectrometer. In this case the muon reconstruction starts from an Inner Detector track that is further extrapolated to the Muon Spectrometer and combined with either a track segment on the middle or lower Muon Spectrometer station or to an energy deposit in the calorimeter matching a muon signature; this way one obtains a "Segment tagged" (ST) muon.

In the reconstruction process ID tracks and MS segments that have been previously associated to a combined muon are not used to form a tagged or a standalone muon so that no overlap occurs.

Finally, last algorithm provides the so-called "Calorimeter" (Calo) muons. As for the Segment tagged, the reconstruction starts from an ID track matched to a calorimeter energy deposit compatible with a muon.

The reconstruction efficiency for both reconstruction families has been evaluated by using muons coming from the Z boson candidates found in 2010 ATLAS data [64]: some examples of the performance achieved as a function of the pseudorapidity and the transverse momentum are reported, respectively, in figures 3.16 and 3.17. In particular, these plots shows the good performance of the CB algorithms and the further improvement obtained when they are complemented by ST algorithms.

3.2 The ATLAS trigger System

The aim of the Trigger System is to identify and select for the storage the most interesting physics events among a huge amount of background.

In order to accomplish these tasks the ATLAS trigger has been designed to select events in a "inclusive" way and to be, at the same time, robust, flexible and redundant. In particular, the specific goals of Trigger System are:

- to select particles featuring events of interest, like charged leptons in different ranges of transverse momentum;

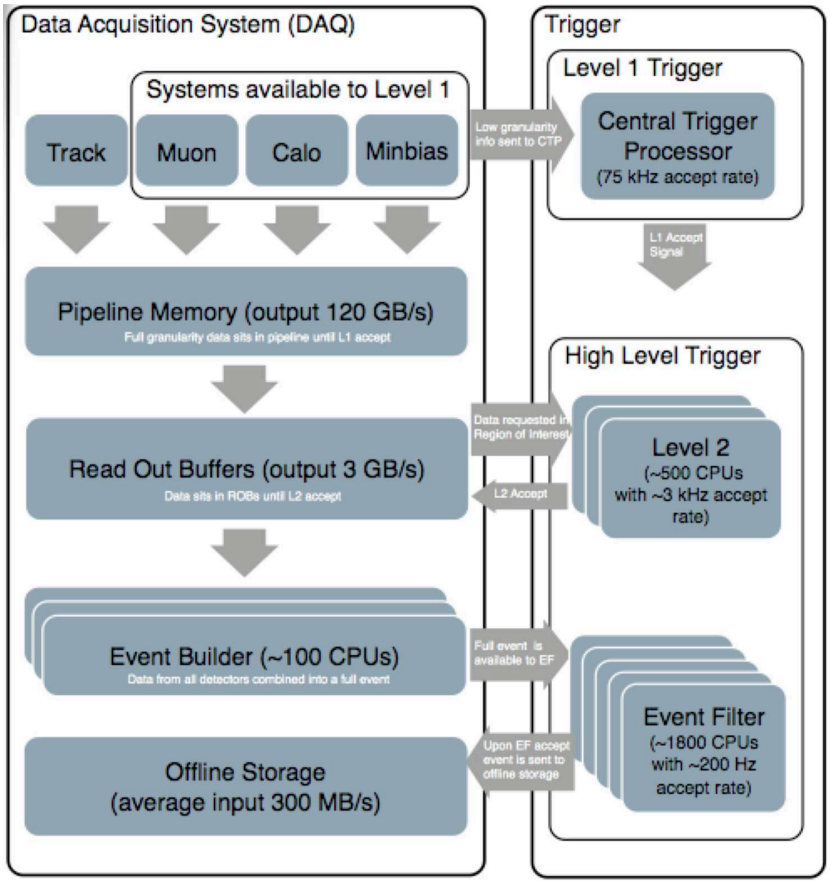


Figure 3.18: ATLAS Trigger System block diagram.

- to select energetic jets;
- to select missing transverse energy;
- to provide the largest possible coverage;
- to guarantee good resolution and efficiency;
- to perform bunch crossing identification, since the offline event building rely on the trigger information to distinguish events.

In order to cope with these requirements ATLAS Trigger uses fast detector systems and is layered in three levels, aiming at progressively reducing the LHC collision rate of $\mathcal{O}$(GHz) to $\sim$ 300 Hz.

- **The Level-1 (L1)** is the first level of the trigger system and is hardware based. The systems providing information to this trigger are the

Calorimeters, and the Muon Spectrometer. At this stage a first identification of interesting trigger objects like jets, photons, electrons and muons, etc., in a defined "Region of Interest" (RoI) is performed. If these objects accomplish specific requirements (thresholds) the event is accepted ("L1 accept") and the information is buffered in the Read-Out System (ROS) waiting for the second trigger level's decision. The rate is so reduced to $\sim$ 75 kHz with a latency of $\sim$ 2.5 μs, where about 1 μs is accounted for the cable-propagation delays alone.

- **The Level-2 (L2)** is the second level and is software based. In a first step it analyses the L1 RoIs and eventually confirms them. Then, by adding the information of the Inner Detector tracks it builds "particle" objects. L2 operates a further reduction of the rate up to $\sim$ 3 kHz with a latency of $\sim$ 10 ms.

- **The Event Filter (EF)** is the last level of event selection and performs the final rate reduction up to $\sim$ 200 Hz with a latency of 1 s. It uses the information coming from the various detectors and performs an on-line reconstruction of all the objects in the events by means of the offline reconstruction algorithms adapted to the online environment.

Being both software-based, L2 and EF usually are common denoted as High Level Trigger (HLT).

The ATLAS Trigger architecture, schematically sketched in figure 3.18, applies to all L1 detector systems (calorimeters, muons, minimum bias [12]). Next subsection is dedicated to the Muon Trigger since it is the one used in this analysis, while next section will briefly explain the steps that starting from the on-line event selection lead to the final objects used in the analysis.

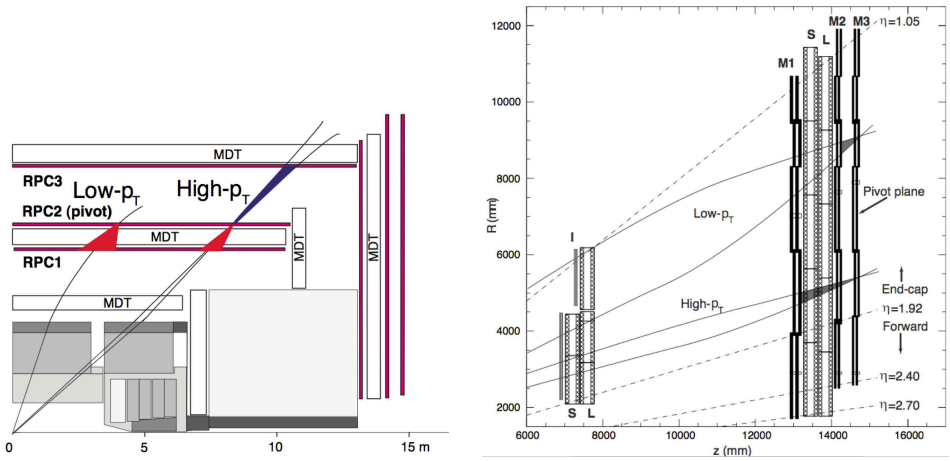
3.2.1 The Muon Trigger

The Muon Trigger System has been designed in order to fulfill the general requirements of the ATLAS Trigger System. More in particular, it has to

1. be able to select events of interest containing both low-p_T (like in B-Physics) and high-p_T objects;

2. work in a high background environment;

[12] The minimum bias trigger is provided at L1 by the Minimum Bias Trigger Scintillators (MBTS): this system is so called because it's placed in the two forward regions and is capable to trigger the collision events without any requirement on the particles involved, having the minimum possible bias on the trigger.



(a) L1 barrel logic (b) L1 endcaps logic

Figure 3.19: L1 muon trigger layout.

3. identify the bunch crossing;

4. provide a trigger decision with a good resolution;

5. measure roughly the muon transverse momentum on a time scale of $<$ 2 μs.

6. work correctly also with the cavern background which can affect performance.

The L1 Muon Trigger makes use of RPC and TGC chambers. It selects muon tracks coming from the IP having a transverse momentum above a given threshold; usually it has two working regimes, classified as "High" or "Low-p_T".

The L1 trigger logic is schematically shown in figure 3.19, for both barrel (left) and endcaps (right).

The lower trigger thresholds in the barrel involve the inner and the middle stations: the former is called "Low-p_T confirmation plane", the latter "Pivot" plane. The presence of hits on the Pivot plane instantiates the Low-p_T algorithm if in both the η and ϕ projections a coincidence in three out of four strip planes is found [13]. Then, hits are searched on the Low-p_T confirmation plane within a geometrical muon "road". The road center is defined by the

[13]This trigger logic is usually referred as "majority logic". This logic, together with the use of the two projections, are dictated by the aim to reduce the rate of "accidental" coincidences.

line of conjunction of the hits on these two planes up to the Interaction Point, corresponding to an infinite momentum track, while the width depends on the p_T threshold. Low-p_T seeds the High-p_T trigger: if a coincidence of hits is found also in the outer station, the "High-p_T confirmation plane", then the high p_T thresholds are fired.

This logic of subsequent seeding starting from the Pivot plane holds also in the endcaps. This time the Pivot plane is the outermost from the IP: the Low-p_T trigger is fired if there is a coincidence of three out of four planes in the two outermost TGC layers, while the High-p_T threshold requires additionally a two out of three coincidence in the bending plane of the triplet of the innermost TGC chambers plus one out of two coincidence in the azimuthal strip planes.

The first estimation of the muon track parameters (η, ϕ, p_T), the RoI[14], is used onwards.

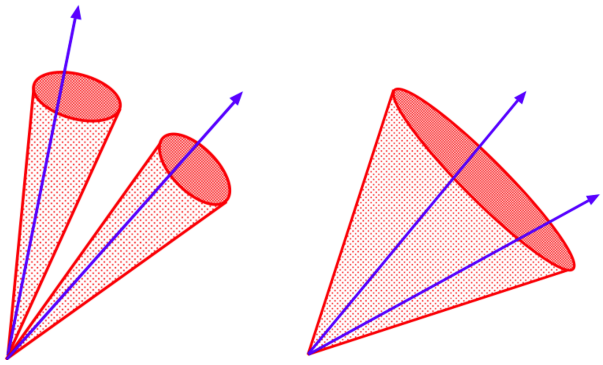
The L1 triggers seed the L2, where different RoIs are processed in parallel. Several algorithms are available, and the processing time increases as the information used: the faster builds a muon track in the MS and extrapolate it to the Interaction Point, while the others achieve better resolution by adding the ID tracks and Calorimeters information. The EF uses the full event data available to select muons by means of offline reconstruction algorithms adapted to the on-line environment. Two main reconstruction strategies are available: *inside-out*, where an ID track is extrapolated to the MS; *outside-in*, where a MS track is back-extrapolated to the IP. As in L2, a further algorithm combines the MS and the ID tracks.

The geometrical coverage of the Muon Trigger in the barrel is about 80% because of the detector feet, the rails carrying the calorimeter supports and the holes for the magnets coils ("ribs"); moreover, there is a crack in the central $(r - \phi)$ plane at $\eta = 0$, for the passage of the cables and services of the Inner Detector, the Central Solenoid and the calorimeters; in the endcaps, instead, it reaches 96%.

Di-muon triggers and performance

All the ATLAS analysis rely on the Trigger capabilities to select events of interest; in particular, for the resonances involving two muons in the final states some sets of di-muon algorithms operating the HLT stage have been provided. This choice is dictated by the limited bandwidth that does not allow to use always single-object triggers.

[14] In the barrel the RoI size is about $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. In the endcaps, instead, is about 0.03×0.03.



(a) Topological trigger (b) Di-muon trigger

Figure 3.20: Di-muon trigger strategies

For the B-Physics studies two different approaches have been implemented: in the first the L2 algorithm is seeded by a di-muon trigger at L1, which produced 2 RoIs, while in the second approach the L2 algorithm starts from a single RoI.

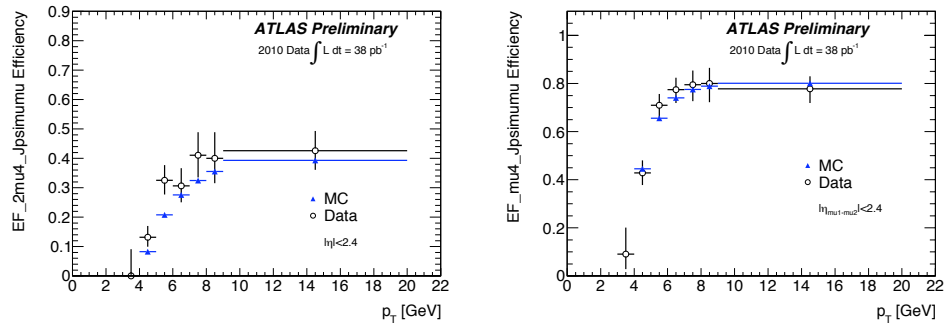
The first approach is denoted as "topological" di-muon trigger: each L1 RoI is separately confirmed at L2, then the two muon candidates are combined into a resonance and a mass cut is applied; finally this chain is confirmed by EF.

In the second approach the L1 RoI is enlarged and in this "Extended RoI" [15] a second muon is searched starting from a Inner Detector track and extrapolating it to the Muon Spectrometer. This second method is particularly useful in the case in which one of the two muons of the resonance is a low-p_T one that may not cross the entire spectrometer or fire the trigger, and in the case in which the angular distance between the tracks of the two muons emitted is smaller than the cone opening angle; this is, as an example, for the J/ψ in figure 3.21 (b).

The first approach, instead, can be more useful in the case in which the two muons have similar energies and the angular distance between the tracks of the two muons emitted is bigger than the cone opening angle, as for the B mesons.

The two approaches are pictorially reported in figure 3.20 while an example of the trigger efficiencies evaluated for the two approaches is reported in figure 3.21: here both plots show the efficiencies evaluated with respect to the offline reconstruction on the $J/\psi \to \mu\mu$ candidates found in 2010 data

[15] The typical Extended RoI dimensions are $\Delta\eta \times \Delta\phi = 0.75 \times 0.75$



(a) Topological efficiency (b) Di-muon efficiency

Figure 3.21: Di-muon triggers efficiencies evaluated on $J\psi \to \mu\mu$ candidates in 2010 data compared to MC as a function of the higher muon transverse momentum.

compared to Monte Carlo simulation. In particular, the plot on the left shows the efficiency as a function of the p_T of the muon with the higher momentum for the topological trigger chain: two muons passing the 4 GeV trigger threshold have been identified at L1 and further confirmed at the HLT in a defined invariant mass region containing the J/ψ mass. Left plot, instead, shows the analogue efficiency for the di-muon trigger threshold.

3.3 The ATLAS Computing Model

The large dimension of the ATLAS experiment and the complexity of the events requiring a huge granularity of the detector systems in order to identify interesting events in a large background environment reflects into a huge event size to be stored. In particular, at nominal LHC luminosity, the ATLAS detector could produce 1.6 MB/event $\times$100 kHz $\sim$ 156 GB/s of data, that are reduced to 1.6 MB/event $\times$ 300 Hz = 320 MB/s thanks to the Trigger System; this corresponds to a huge amount ($\sim$ 2 PB/year) to be stored and analyzed. To cope with these requirements an efficient and distributed computing system has been set-up and its performance are continuously monitored. In the next subsections a short description of this system will be given. In particular, the process that allow to offline convert the trigger output into a data format suitable for the analysis will be illustrated. Moreover, the world-wide distributed ATLAS Computing System [72] will be described.

3.3.1 The ATLAS Data Model

Collisions events filtered by the Trigger and Data Acquisition system are collected in the so-called "data streams". Each stream contains events characterized by common features individuated by the HLT. Data are separated into luminosity blocks corresponding to a few minutes of data taking; this separation allows for the flagging of them into quality type depending on the detector performance.

The ATLAS Event Data Model (EDM) contains different data formats; the output of the trigger system consists of binary "Raw" data, having a mean size per event of $\sim$ 1.6 MB; this size is progressively lowered to a few kB trough the format conversion and skimming of the information.

This bit-format is further converted into an high-level information format called ESD (Event Summary Data). ESDs contain the full reconstruction output in "POOL/ROOT" objects [66], [67] [16] like electrons, jets, muons and the corresponding parameters (η, p_T, etc.). At this stage the nominal event size is of $\sim$ 1 MB.

ESD are then converted into a smaller ($\sim$ 200 kB) format, the AOD (Analysis Object Data) format. This contains a summary of the event reconstruction in the same POOL/ROOT object structure.

AOD files are the common starting point to many analysis; the information contained can be further reduced leading to the DPD (Derived Physics Data), having different size depending on the different selections applied.

Finally, in the ATLAS EDM there are the TAG files: these are database (or ROOT files) used to quickly select events among the AOD and/or ESD files.

The same output formats hold for Monte Carlo events: the passage of particles deriving from generation output through the detector is operated by GEANT4 [68], a software that simulates the detector response of ATLAS and write out the simulated data in the same format of the real detector output.

3.3.2 The ATLAS Distributed Computing Model

Every year of data-taking the ATLAS experiment collect a large amount of data and approximately the same amount derives by the large-scale processing, reprocessing and analysis activities. Moreover, the computing power

[16] ROOT is an object-oriented software framework developed at CERN and commonly used by the high-energy physics community with its own data format. POOL is one of these possible data formats.

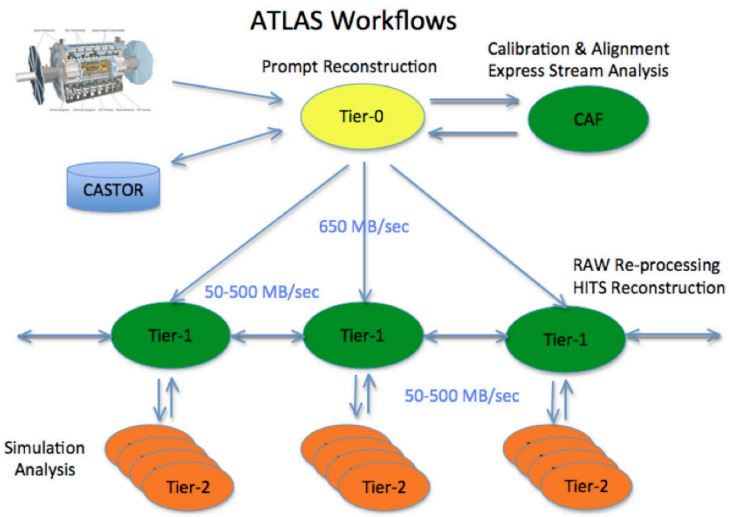


Figure 3.22: The ATLAS workflow in the *Tier* structure.

required to process this huge amount of data is of the order of 10^5 CPUs full time.

The impossibility of concentrating the needed computing power and storage capacity in a single place required the development of a world-wide distributed computing system, which allows efficient data access and makes use of all available computing resources. Thus ATLAS Collaboration embraced a computing model based on the Grid paradigm and on a high degree of decentralization and sharing of computing resources.

The main difference with the previous CERN-based Computing Models is that off-site facilities play a vital role to the operation of ATLAS. In particular, since 2002 a complex set of tools and distributed services, enabling the automatic distribution and processing of the large amounts of data, has been developed and deployed by ATLAS in cooperation with LHC Computing Grid (LCG) Project [73] and with the middleware providers of the three large Grid infrastructures: EGEE [74] in most of Europe and the rest of the world, NorduGrid [75] in Scandinavian countries and OSG [76] in the United States. The Computing Model designed by ATLAS is a hierarchical structure organized in different *Tiers* (see figure 3.22). The *Tier0* facility is based at CERN and is responsible for first-pass processing and archiving of the primary Raw data and their distribution to the *Tier1s*. The ten *Tier1* facilities word-wide distributed have to store and guarantee a long-term access to Raw and derived data and provide all the reprocessing activities. Each *Tier1* heads up a set of *Tier2* sites grouped in regional Clouds. *Tier2s* are

the centers designed for the users analysis and provide all the Monte Carlo simulation capability.

In order to access grid resources, ATLAS members are organized in the *Atlas Virtual Organization* (ATLAS VO); the computing information and the privileges of all ATLAS members are stored in a VOMS database.

Data distribution takes place through the Distributed Data Management (DDM) system, which provides data replication in the Collaboration sites by using the different data transfer protocols in a transparent way and keeps tracks of data distribution in Data Catalogs.

The ATLAS Production System, ProdSys/Panda, is a central system which schedules all organized data processing and simulation activities. The production System makes use of the computing resources enabled by the three different middlewares (LCG/EGEE, NDGF/Nordu-Grid and OSG/Panda) joining ATLAS VO.

The ATLAS software is world-wide distributed by means of the ATLAS Software Installation System for LCG/EGEE [78], responsible for access, monitor, install, test e publish information about all resources made available to ATLAS Collaboration by LCG.

Finally, a set of Distributed Analysis Tools is provided by the Collaboration in order to provide transparent command line or graphical interface between the user and the grid infrastructure, hiding grid technicalities to users.

Chapter 4

The $B_s \to \mu^+\mu^-$ search with ATLAS

In this chapter the analysis performed on the $B_s \to \mu^+\mu^-$ decay will be described. The strategy adopted is not to measure directly the $B_s \to \mu^+\mu^-$ branching ratio, but to obtain it from the branching ratio of another well measured B meson decay taken as reference which features almost the same kinematics as the signal; this choice allow to cancel out the main sources of systematic uncertainties. As shown in this chapter, the evaluation of the branching ratio can be reduced to the evaluation of yield of the reference channel and of the relative acceptance and efficiencies. The analysis has been performed by using 2010 ATLAS data corresponding to an integrated luminosity of $\sim 40\ pb^{-1}$ and it is considered as baseline analysis for the 2011 data which is going to be finalized soon.

4.1 The analysis strategy

The number of observed events N_k^{obs} in the detector for a given particle decaying into the channel k is related to the number of events expected by the expression:

$$N_k^{obs} = \alpha_k \cdot \epsilon_k \cdot \sigma^{MC} \cdot Br(k)\mathcal{L} \tag{4.1}$$

where α_k accounts for the geometric and kinematic acceptance, ϵ_k includes the trigger, the reconstruction and the selection efficiency, $\mathcal{L}$ is the integrated luminosity, σ is the cross section for the particle's production and $Br(k)$ is the branching ratio. In the case of the rare decays generally the systematic uncertainties dominate this measurement as, apart form the integrated luminosity, all the quantity entering in equation 4.1 are evaluated using a Monte Carlo simulation of the process and detector response.

This necessity to reduce as much as possible the systematic uncertainties on the branching ratio measurement of the $B_s \to \mu^+\mu^-$ decay channel lead to perform this measurement by normalizing it to the branching ratio of another B meson decay channel taken as reference, so that in the normalization most of the systematics cancel out and no direct measurement of the luminosity and of the cross-section of the process is needed.

In this view, the expression of the branching ratio reads

$$Br(B_s \to \mu^+\mu^-) = \frac{N_{\mu\mu}}{N_{ref}} \times \frac{\mathcal{L}_{ref}}{\mathcal{L}_{\mu\mu}} \times \frac{\epsilon_{ref}\alpha_{ref}}{\epsilon_{\mu\mu}\alpha_{\mu\mu}} \times \frac{f_{ref}}{f_s} \times Br(ref) \qquad (4.2)$$

where the ratio $\frac{f_{ref}}{f_s}$ takes into account the difference in the b-quark fragmentation probabilities, which is directly connected to the ratio of the relative cross sections.

The choice of the reference channel is leaded by two main reasons:

1. the decaying meson should have quite the same kinematics as the B_s;
2. the branching ratio should be as high as possible in order to have a sample of events abundant enough to reduce as much as possible the statistic uncertainties entering in the measurement.

These reasons lead to choose for this analysis as reference channel the $\mathrm{B}^\pm \to J/\psi K^\pm$ decay channel, where the J/ψ further decay in two muons having opposite charge; equation 4.2 can be so re-expressed as follows

$$Br(B_s \to \mu^+\mu^-) = \frac{N_{\mu\mu}}{N_{B^\pm}} \times \frac{\epsilon_{B^\pm}\alpha_{B^\pm}}{\epsilon_{\mu\mu}\alpha_{\mu\mu}} \times \frac{1}{\epsilon_N}\frac{f_u}{f_s} \times Br(B^\pm \to J/\psi K^\pm) \times Br(J/\psi \to \mu^+\mu^-) \qquad (4.3)$$

where ϵ_N accounts separately for the optimization performed on the signal variables in order to enhance the power of discrimination from background (see section 4.1.1) and the luminosity factors cancel out since same integrated luminosity is used. Also the contribution to the efficiency from the trigger cancels out because the two trigger chains used were running the same algorithm, imposing only a different mass window cut, and have run un-prescaled all along the data taking period.

Equation 4.3, when setting $N_{\mu\mu} = 1$, provides the so-called *Single Event Sensitivity* (SES)

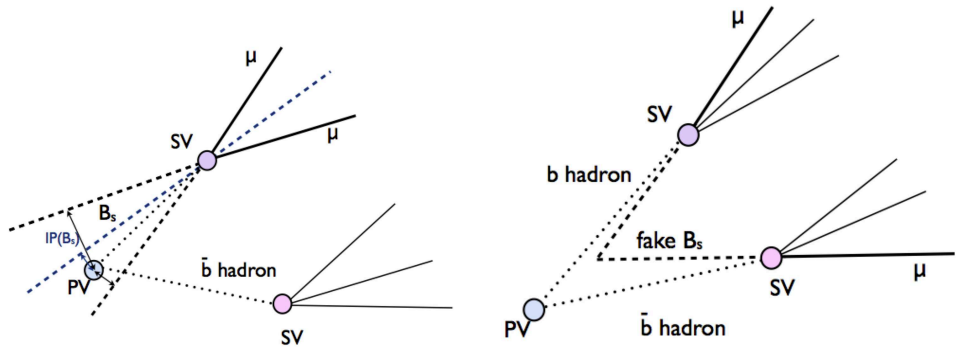
$$SES = \frac{1}{N_{B^\pm}} \times \frac{\epsilon_{B^\pm}\alpha_{B^\pm}}{\epsilon_{\mu\mu}\alpha_{\mu\mu}} \times \frac{1}{\epsilon_N}\frac{f_u}{f_s} \times Br(B^\pm \to J/\psi K^\pm)Br(J/\psi \to \mu^+\mu^-) \qquad (4.4)$$

This is an important estimator of the sensitivity of the experiment for the measurement because it is uncorrelated to the actual number of $N_{\mu\mu}$ signal events but contains all the quantities estimated from the data analysis and their relative errors. On the other hand the number of the candidates B_s found after the selection in a given mass window, compared to the expected background yield in the same mass window, will set the real limit.

The ATLAS B Physics Rare Decays group has adopted the blind analysis strategy, meaning that all the quantity entering the SES have been evaluated by excluding in the data sample the signal region, corresponding to a mass window of $\pm 300 MeV$ around the B_s invariant mass (5.0663 ± 0.0006 GeV), until the analysis was considered finalized and approved by the Collaboration. This means that, in general, Monte Carlo samples were used as reference for the signal and for the peaking background, while to model the non-resonant background the data belonging to the sidebands of the signal region were used (see 4.7 for details). In truth, the events in the sidebands have been used both for the cut optimization procedure and for the estimation of the number of background events in the signal region. In order to avoid biases, as shown later 4.4.6, the sidebands data have been split in 2 sample: the odd-numbered events were used in the optimization of the cuts, while the even numbered events for the estimation of the background.

This study has been entirely based only on 2010 ATLAS data, corresponding to an integrated luminosity of $\sim$40 pb^{-1} and on "MC10" Monte Carlo samples (see subsection 4.3.1) because the analysis based on 2011 data is in the final stage of the approval from the Collaboration so data are still blinded and then we cannot quote here the limit. This analysis is the baseline used also for the 2011 data even if the much higher statistics makes different the evaluation and the contribution of the different source of uncertainty.

In the SES, the reference yield $N_{B^\pm}$ is measured by using data as will be described in section 4.2, while the product $\epsilon \times \alpha$ is evaluated by using Monte Carlo samples for both signal and reference channel; ϵ_N is evaluated by employing both data (as background) and Monte Carlo (signal) in the optimization procedure that will be illustrated in section 4.4.3, while the remainder quantities are taken from other measurements.
It is important to note that the major sources of systematic uncertainties in equation 4.3 come from the $N_{\mu\mu}$ extraction and from the ratio $\frac{f_u}{f_s}$, actually measured with an uncertainty of $\sim$10% by combining the Particle Data Group [69] result on the $B^\pm \to J/\psi K^\pm$ branching ratio and latest LHCb [70] result on the b-hadron production fractions; the former, together with the upper limit evaluation will be described separately in the next chapter.



(a) Signal (b) Background

Figure 4.1: Left: signal decay topology. Right background decay topology from semileptonic B meson decays ($B \to \mu^+ X$, $\bar{B} \to \mu^- Y$).

4.1.1 Signal channel features

The signal event topology results quite simple; in fact, given its long lifetime ($\tau = 1.47 \pm 0.03\, ps$, $c\tau = 441\ \mu m$), after the production at the primary vertex (PV) the B_s meson travels in the detector for an appreciable distance before decaying. The muons emitted result isolated and the sum of their transverse momenta lies at a small angle with respect to the B_s flying direction. These features, together with the invariant mass of the two muons, allow to characterize signal sample with respect to the background. In fact, the main background is expected to originate from true di-muon events, dominated by (prompt) Drell-Yan pairs ($pp \to \mu^+\mu^-$) and non-prompt heavy flavor semileptonic decays ($pp \to b\bar{b} \to \mu^+\mu^- X$); in the first case the muons are emitted in the PV, so the secondary vertex and all quantities that can be built by using this information are useful tools to detect it. The second case can be detected also by using the muons isolation and the di-muon direction. The event topology is schematically sketched in figure 4.1 for both signal (left) and non-prompt heavy flavor semileptonic decays (right). More details on the separation variables will follow.

4.2 The $B^\pm$ reference channel

$B^\pm$ mesons arising from $b\,\bar{b}$ pairs are reconstructed in ATLAS as $J/\psi \to \mu^+\mu^- K^\pm$ decay channel.

Its relatively high branching ratio ($\sim 0.6\%$) allows to use it for many purposes; for instance, it is used as a reference for high-precision B-physics

measurements (like in this analysis), as a calibration tool for flavor tagging in CP violation studies, and, more in general, as an instrument to evaluate the Inner Detector performance. One of the first step of the measurement is the determination of the number of event expected for the reference channel, usually called yield, which depends on the selection applied.

4.2.1 The reference channel selection

The reference yield has been evaluated by using $\sqrt{7}$ TeV $p-p$ collision data taken during 2010, i.e. when the detector was declared to be fully operational and LHC beam stable. The statistics corresponds to an integrated luminosity of $\sim 40\ pb^{-1}$.

Events were selected by the lowest non-topological di-muon trigger threshold within the J/ψ mass window (cfr. sec. 3.2.1).

Additional selection cuts, here following, are applied to clean-up the sample and remove as much as possible background events that will bias our estimation of the yield:

- **Collision Selection**: to veto cosmic muons, events were required to contain at least one reconstructed primary vertex with at least three associated Inner Detector tracks;

- **Track quality**: tracks used for the $B^{\pm}$ meson reconstruction were required to have at least one hit in the Pixel detector, six hits in the SCT an one hit in the TRT tracker;

- **Muon Selection**: events were required to contain at least one couple of opposite charged muons, either formed by two Combined muons or by one Combined and one Segment-Tagged muon (cfr. sec. 3.1.6). The Inner Detector tracks associated with the reconstructed muons should have matched the Track quality requirements; moreover, the highest-p_T muon in the pair should have at least a transverse momentum $p_T > 4$ GeV, while the other at least a transverse momentum $p_T > 2.5$ GeV.

- J/ψ **Selection**: the opposite signed di-muon tracks have an invariant mass in the range $[2.915, 3.275]$ GeV and they are successfully fitted to a common vertex with $\chi^2/ndf < 10$.

- **K Selection**: kaon candidate is associated to an Inner Detector track which satisfies the Track quality criteria and has a transverse momentum $p_T > 2.5$ GeV.

- $B^{\pm}$ **Selection**: J/ψ tracks pairs and kaon track passing the selection above described are considered to form a $B^{\pm}$ candidate if they are successfully fitted to a common vertex with $\chi^2/ndf < 6$ and if their invariant mass lies in the range $4 \leq \mathrm{M_B} \leq 7$ GeV. Moreover, $B^{\pm}$ candidates are required to be reconstructed within the pseudo-rapidity range $|\eta| \leq 2.5$ and to have a transverse decay length $\mathrm{L_{xy}} > 0.3$ mm and a transverse momentum of $p_T > 6.0$ GeV. The transverse decay lenght is expressed in terms of the refitted $B^{\pm}$ meson transverse momentum and of the vector pointing from the primary vertex to the B vertex (both computed in the transverse plane), $\vec{X}_{PV}$, as

$$L_{xy} = \vec{p}_T \cdot \vec{X}_{PV}. \tag{4.5}$$

4.2.2 The reference yield extraction

The $B^{\pm}$ yield is estimated using an un-binned maximum likelihood fit in the invariant mass range [4930,5630] MeV.

In this mass range the main source of background is given by the combinatorial background coming from the process $b\bar{b} \to \mu^+\mu^- X$. Other significant contributions may come from partially reconstructed decays like $B^{\pm 0} \to K^{*\pm 0} J/\psi$ or $B^+ \to \chi_{c1,2}$ (in the mass region on the left of the $B^{\pm}$ peak) and from the $B^{\pm} \to J/\psi\pi^{\pm}$ (in the mass region on the right of $B^{\pm}$ peak) when the pion is misidentified as kaon; unfortunately there is no particle identification on the third track used for the $B^{\pm}$ vertex, so the latter decay channel can not be separated from the $B^{\pm} \to J/\psi K^{\pm}$ decay. The fraction of the $B^{\pm} \to J/\psi\pi^{\pm}$ background relative to the reference channel is $\sim 4.8\%$: this appeared to be totally negligible with respect to the statistical uncertainties and therefore not accounted for in the fit model.

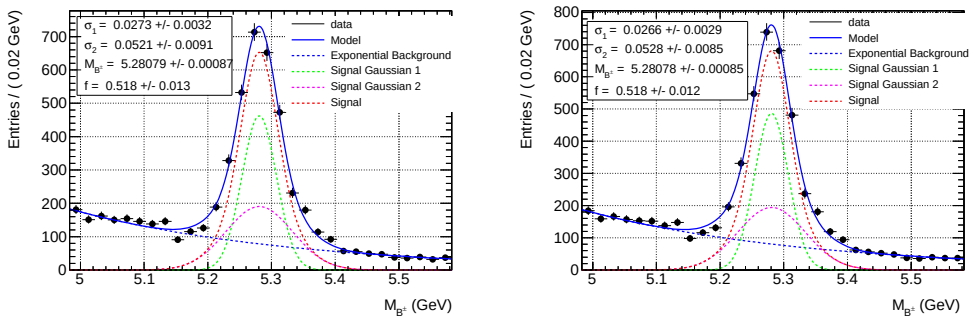
To extract the yield a maximum likelihood fit of the mass region is performed with a likelihood function defined as a composite probability density function (PDF)

$$\mathcal{L} = f \cdot S(M_{J/\psi K^{\pm}}) + (1 - f) \cdot B(M_{J/\psi K^{\pm}}) \tag{4.6}$$

where $S(M_{J/\psi K^{\pm}})$ is the signal model, $B(M_{J/\psi K^{\pm}})$ the background model and f the fraction of signal events in the sample. Note that in this model the composite PDF is normalized, providing that both $S(M_{J/\psi K^{\pm}})$ and $B(M_{J/\psi K^{\pm}})$ are normalized; moreover, each PDF can be on its turn a composite PDF.

In particular, to model the combinatorial background an exponential PDF is assumed

$$B(M_{J/\psi K^{\pm}}) = \frac{1}{N_B} e^{-\lambda M_{J/\psi K^{\pm}}}, \tag{4.7}$$



(a) STACO (b) MUID

Figure 4.2: Fit on the $B^{\pm}$ invariant mass spectrum for the yield determination using the both STACO (a) and MUID (b) muon reconstruction families. The solid blue curve is the fit overlaid on top of binned data distribution (black circles); the dotted curves correspond to the various fit model components.

where N_B is a normalization factor and λ a free parameter, while for the signal a composite PDF made up of two gaussians $G_{1,2}(M_{J/\psi K^{\pm}})$ constrained to have the same mean $M_{B^{\pm}}$ but free standard deviation $\sigma_{1,2}$ has been used:

$$S(M_{J/\psi K^{\pm}}) = \frac{1}{N_S}\left(aG_1(M_{J/\psi K^{\pm}}) + (1-a)G_2(M_{J/\psi K^{\pm}})\right) . \tag{4.8}$$

Here N_S is the normalization factor, a the fraction of the event in the first gaussian and

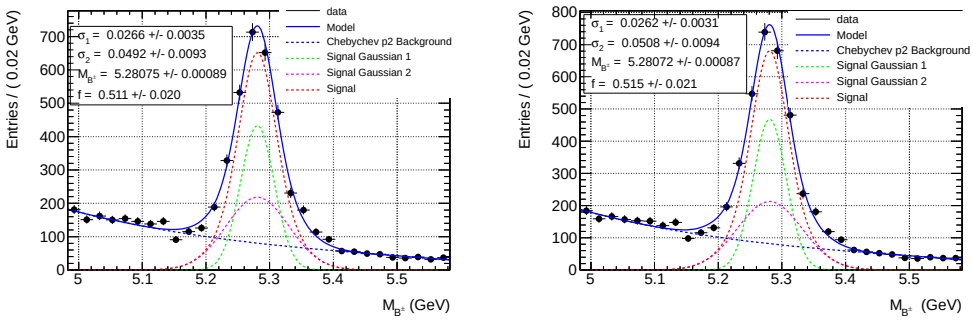
$$G_{1,2}(M_{J/\psi K^{\pm}}) = \frac{1}{\sqrt{2}\sigma_{1,2}} e^{-\frac{(M_{J/\psi K^{\pm}} - M_{B^{\pm}})^2}{2\sigma_{1,2}^2}} . \tag{4.9}$$

The two different widths in the gaussians aim to account for different resolution effects on the $B^{\pm}$ invariant mass.

The maximum likelihood fitting technique provides the five free parameters of the fit, i.e. f, λ, $M_{B^{\pm}}$ and $\sigma_{1,2}$ together with their errors and the covariance matrix. These parameters allow to estimate the number of signal and background events integral of the PDFs in the chosen mass window.

The fit has been performed separately by using muons from the two different reconstruction families, STACO and MUID (see section 3.1.6): the results of the fits are shown in figure 4.3 and the number of estimated $B^{\pm}$ candidates is $N^{S,E}_{B^{\pm}} = 2754 \pm 54$, with a χ^2/ndf equal to 1.62 for STACO, and $N^{S,E}_{B^{\pm}} = 2834 \pm 50$ with χ^2/ndf=1.37 for MUID.

One of the possible source of systematic uncertainty on these results is the fit model used.



(a) STACO (b) MUID

Figure 4.3: Fit on the $B^{\pm}$ invariant mass spectrum for the systematic uncertainties evaluation on yield determination using the both STACO (a) and MUID (b) muon reconstruction families. The solid blue curve is the fit overlaid on top of binned data distribution (black circles); the dotted curves correspond to the various fit model components.

In order to estimate it the combinatorial background in the fit model reported in equations 4.6 and 4.7 has been changed from an exponential to a second order Chebychev polynomial PDF, i.e.

$$B(M_{J/\psi K^{\pm}}) = \frac{1}{N_B}(c_0 T_0(M_{J/\psi K^{\pm}}) + c_1 T_1(M_{J/\psi K^{\pm}}) + c_2 T_2(M_{J/\psi K^{\pm}})) \quad (4.10)$$

where T_i is a i^{th} order Chebychev polynomial. In this case, the number of estimated number of events for the STACO family is $N^{S,C}_{B^{\pm}} = 2732 \pm 57$, while for MUID $N^{M,C}_{B^{\pm}} = 2823 \pm 53$; the corresponding χ^2/ndf are equal to 1.65 and 1.37 respectively.

The systematic error for each reconstruction family as been computed as

$$\frac{\delta N_{B^{\pm}}}{N_{B^{\pm}}} = \frac{N^{E}_{B^{\pm}} - N^{C}_{B^{\pm}}}{N^{E}_{B^{\pm}}} \quad (4.11)$$

Thus, in the case of the STACO family it is equal to 0.8%, while for the MUID family it goes down to 0.4% .

4.3 Monte Carlo tuning

Simulated events for both reference and signal decay channel are used in this analysis primarily for the determination of efficiencies and acceptance, as they enter in formula 4.3; in particular, being the signal a rare process for which there is no experimental evidence up to now that may confirm the SM

branching ratio predictions, particularly caution should be used in trusting Monte Carlo simulation.

More in details, the use of simulation introduces different potential sources of uncertainty:

- limited Monte Carlo statistics,
- inconsistency of the physics model contained in the Monte Carlo with respect to the data,
- inaccuracy and incompleteness in the modeling of the detector.

While a larger generation can take care of the first source, there is no immediate solution to the modeling problems.

Additional issues may come if kinematic selections on the phase space of the generated particles are applied, so that acceptance evaluation may be compromised.

During the analysis work the official Monte Carlo samples was found to have, unexpectedly, some cuts at generator level on the final state particles which propagate to the physical quantities distributions used in the analysis making them quite different for B_s and the B^+.

This fact is in contrast with the initial hypothesis that the signal and the reference channel have quite the same kinematics, so the choice of expressing the branching ratio for the signal as relative to the reference channel is no longer convenient.

In order to evaluate the bias induced by these cuts and try to correct it some new Monte Carlo samples have been generated by removing the cuts on the final states: the procedure adopted is illustrated in section 4.3.2.

To correct for the modeling problem, instead, a data-driven approach was used. Signal observed in data were used, in fact, to build per-event weights aimed at re-tuning our Monte Carlo spectra based on the B^+ candidates observed in data; the procedure employed implicitly corrects also for detector performance discrepancies and other potential kinematic biases. The determination of such weights may result affected by the limited signal statistics on data; however, it has the advantage of automatically including also the previously mentioned systematic effects, converting them into a statistical uncertainty. More details about the Monte Carlo samples and their re-tuning will be given later in subsection 4.3.3.

4.3.1 Monte Carlo datasets

Monte Carlo samples are all generated with PYTHIA [79] in conjunction with PYTHIAB [80]. $b\bar{b}$ processes generated are selectively forced to the decays

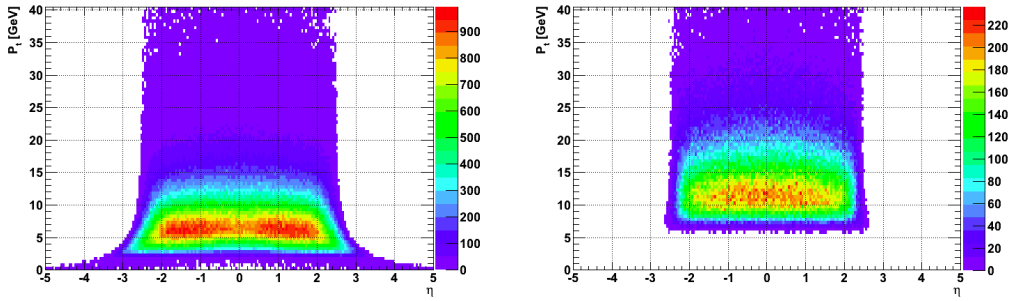
of $B_s \to \mu^+\mu^-$, or $B^+ \to J/\psi K^+$ and $J/\psi \to \mu^+\mu^-$. In order to enhance the sample production efficiency, only events with one b quark within $|\eta| < 2.5$ and $p_T > 2.5$ GeV have been selected, then the hadronization of the b quark was repeated until all requirements are satisfied. In particular, in the official Monte Carlo samples generator level selections have been applied on the Monte Carlo true value of the final state kinematics: the existence of at least one muon with $p_T > 2.5$ GeV and $|\eta| < 2.5$ is required, followed by asking that all the muons in the final state satisfy the same requirements, while all hadrons in the final state (if any) are requested to have $p_T > 0.5$ GeV and $|\eta| < 2.5$. The full ATLAS simulation based on GEANT [68] is then applied to these samples, which are subsequently analyzed with the same tools we use for actual data.

Additional and relatively abundant Monte Carlo samples for both $B_s \to \mu^+\mu^-$ and $B^+ \to J/\psi(\mu^+\mu^-)K^+$ decays have been privately generated by using PYTHIA in conjunction with PYTHIAB as for the official samples, this time accepting only events with one b quark within $|\eta| < 4$ and $p_T > 2.5$ GeV and removing all cuts on the final state particles. Since these samples have been used in the analysis with the only purpose to evaluate the bias induced by the generator level selections on the final state kinematics (see next section), no simulation of the ATLAS detector has been performed on them.

4.3.2 Monte Carlo Generator Level bias and correction

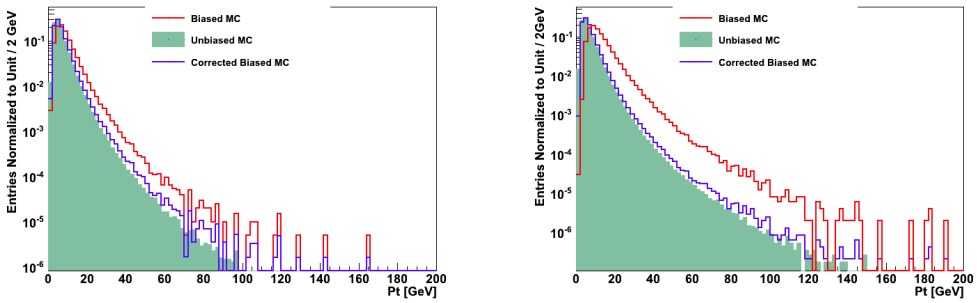
The kinematic requirements at generator level on both signal and reference Monte Carlo samples so far described, although less tight than the analysis cuts, bias the b quark spectra differently between B_s and B^+, being the signal decay chain a two-bodies decay while in the reference chain there are three particles in the final state.

This bias is evident in figure 4.4, where the B_s and B^+ mesons spectra are plotted in the η, p_T variables: in particular, the region between 0 and 7 GeV is almost empty for the B^+ meson, while B_s spectrum shows a discontinuity for $\eta \in [-0.5, 0.5]$and $p_T \in [5, 7]$ GeV not so evident in the case of the B^+. In order to correct for this bias, Monte Carlo events have been generated by removing all cuts on the final state particles for both reference and signal channels: p_T distribution so obtained for both B_s and B^+ mesons is reported in figure 4.5. Both figures confirm that with the "unbiased" generation the spectra of B_s and B^+ mesons are quite the same; to better verify this statement a two-dimensional histogram (fig. 4.6) containing the ratio between the B^+ and the B_s spectra, together with one containing the relative errors on this ratio, has been produced and it confirms that within the uncertainties the two spectra are compatible.



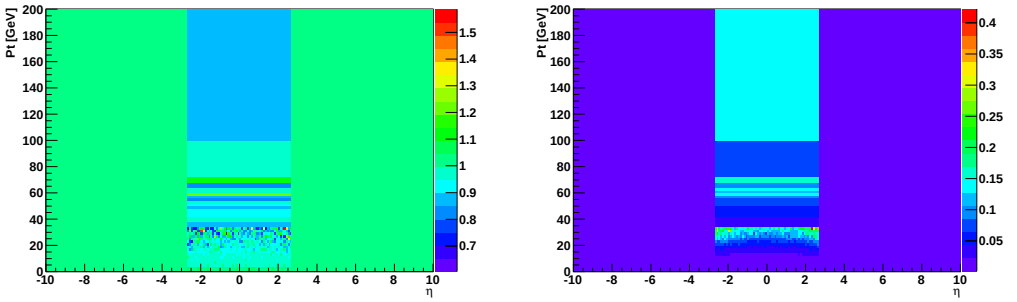
(a) B_s spectrum (b) B^+ spectrum

Figure 4.4: B_s (left) and B^+ truth spectra in the η, p_T plane as in official Monte Carlo samples.



(a) B_s spectrum (b) B^+ spectrum

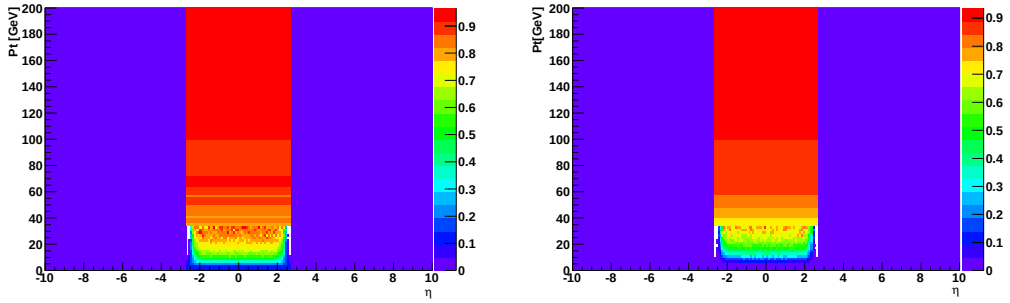
Figure 4.5: B_s (left) and B^+ truth spectra in the η, p_T plane generated by removing cuts on final state particles.



(a) B^+/B_s spectrum (b) Relative errors on B^+/B_s spectrum

Figure 4.6: Ratio of the normalized B^+ and B_s truth spectra (left) and relative errors (right) in the η, p_T plane obtained by using unbiased samples.

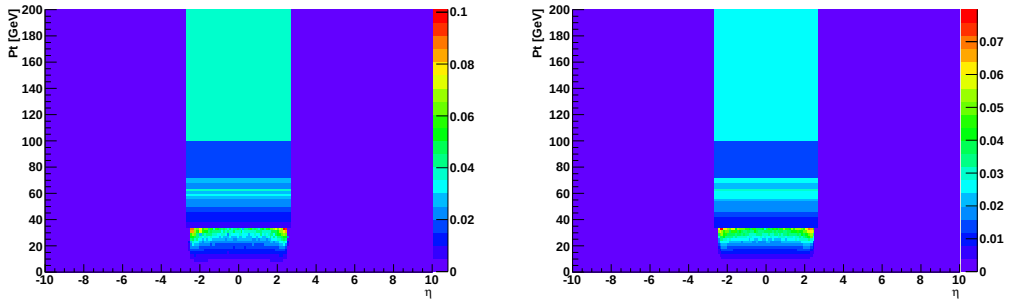
The generated unbiased samples have been then used to build two-dimensional



(a) B_s efficiency (b) B^+ efficiency

Figure 4.7: B_s (left) and B^+ efficiency maps in the η, p_T plane for the generator level selections.

histograms containing correction factors for this bias.
For each decay mode, the efficiencies for the generator level selections have been calculated dividing bin by bin the two-dimensional histograms in the η, p_T plane containing the events which survive the final (generator level) selection cuts (FS) and the ones obtained by using the total number of generated events; thereof, they correspond to the efficiencies for the generator level selections. These efficiencies can be later used to re-weight the event and remove the bias.

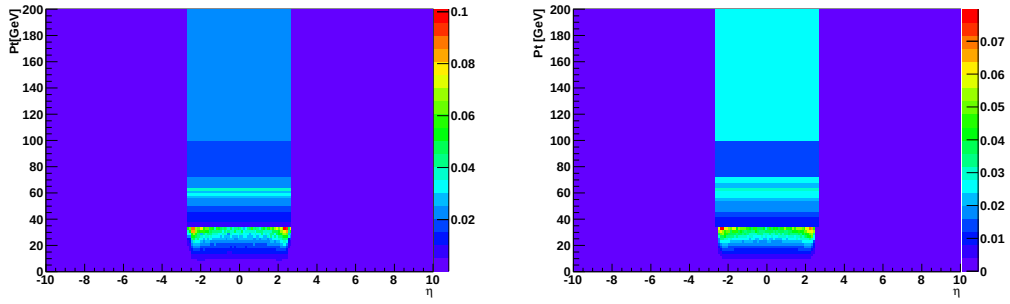


(a) B_s efficiency Upper Errors (b) B^+ efficiency Upper Errors

Figure 4.8: B_s (left) and B^+ efficiency upper error maps in the η, p_T plane for the generator level selections.

Figure 4.7 shows the efficiency maps so built, while the upper and the lower relative errors [1] on these efficiencies are reported, respectively, in fig-

[1]To compute the efficiency uncertainties in the extreme cases where it was close to



(a) B_s efficiency Lower Errors (b) B^+ efficiency Lower Errors

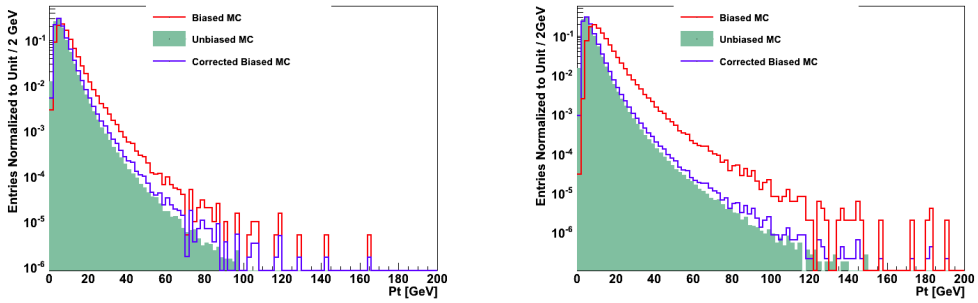
Figure 4.9: B_s (left) and B^+ efficiency lower error maps in the η, p_T plane for the generator level selections.

ures 4.8 and 4.9.

To correct the biased samples, a weight W^{GL} is assigned to each event according to the values of $\bar{\eta}$ and $\bar{p_T}$) of the B meson:

$$W^{GL}\left(\bar{\eta}, \bar{p_T}\right) = 1/\nu_{ij}\,. \tag{4.12}$$

where ν_{ij} is the content of bin corresponding to the $\bar{\eta}$, $\bar{p_T}$ of the efficiency histograms. The weights lie in a limited range of validity of the maps, that is for $|\eta|$<2.5 and p_T <6 GeV of the B meson; outside these ranges or in case ν_{ij} is equal to zero the weight is assigned equal to zero. In order to check



(a) B_s p_T comparison (b) B^+ p_T comparison

Figure 4.10: Comparison between the unbiased Monte Carlo (filled green) and the biased Monte Carlo p_T distributions before (red line) and after (blue line) the corrections for B_s (left) and B^+ (right).

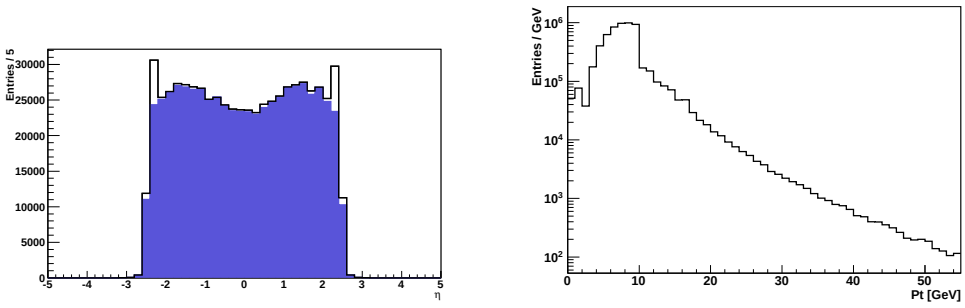
one or zero an approach "á la Feldman Cousin " was used with a confidence level for the intervals of 68% [81].

the validity of these weights, they have been applied to the "biased" Monte Carlo samples and the results have been compared to the newly generated samples: figure 4.10 shows this comparison for the truth p_T distributions of the B_s and the B^+ mesons respectively. In the re-weighted distributions the mean correctly shift to the left towards the mean value of the unbiased distribution; however there is no complete superposition given the limited ranges of validity of the corrections.

The binning of the histograms has been chosen on order to

- be compatible with the histograms used for the data-driven re-weighting described in the following section, as the two corrections will be used simultaneously;
- reduce as much as possible distortion effects on the shapes of the corrected variables due to the finite size of the bin width.

This means that generally these 2D histograms follow the baseline binning choice used for the data-driven weights i. e. a width equal to 2 GeV in p_T and to 0.1 in η, but finest in the low-p_T region and on the η acceptance boundaries ($\pm$ 2.5).



(a) B_s η distribution (b) B^+ $\mathrm{p_T}$ distribution

Figure 4.11: Left:B_s η corrected distributions by using a non optimized (black line) and an optimized (filled violet) binning. Right: "step-like" behavior of the corrected B^+ p_T by using a non-optimized binning.

Some examples of the issue addressed by the latter guideline are shown in figure 4.11. In particular, in the left plot the black line corresponds to the corrected η distribution of the B_s meson by using non-optimized binning for the correction maps: two un-natural peaks are present for η values $\sim \pm 2.5$, while by using an optimized binning (filled violet) this bad behavior disappears. The binning optimization took care also for the "step-like" behavior

of the corrected p_T distributions, illustrated in the right plot of figure 4.11 in the case of the B^+ meson.

4.3.3 Kinematic data-driven re-weighting

Monte Carlo samples have been further corrected in order to take into account discrepancies with the observed data; this correction allows to fold into a statistical uncertainty (deriving from data) all the systematic uncertainties in the Monte Carlo samples deriving from the B meson kinematics and the detector response modeling.

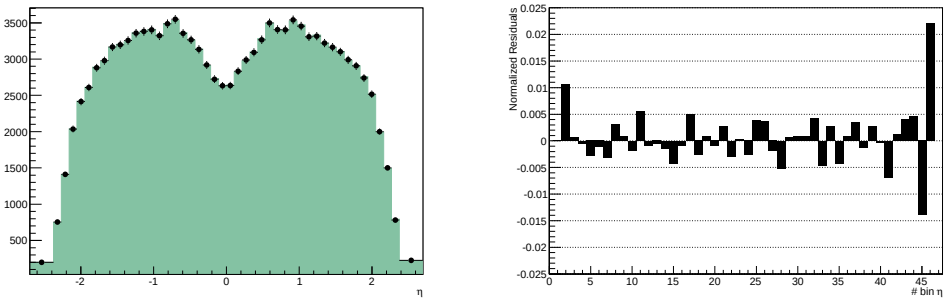
Monte Carlo data-driven event weights have been determined with an iterative method, by comparison of the reconstructed Monte Carlo and data quantities for the same signal. Monte Carlo events have been corrected for the generator level selections as described in the previous section, while for data sideband-subtracted $B^{\pm} \to J/\psi K^{\pm}$ events (where the subtraction is performed as will be described later on) have been used. The limited signal statistics on data led to choose a small number of variables plotted on histograms having a reasonably coarse binning. The most feasible approach resulted into the use of two 1D sets of weights in true p_T and η of the candidate B, determined by the ratio of the normalized 1D p_T and η spectra in data (D) and Monte Carlo (MC). From Monte Carlo (cfr. fig 4.4) it is evident that the correlations between the $\bar{p_T}$ and $\bar{\eta}$ of the B meson can be neglected and then to each event can be assigned a weight

$$W\left(\bar{p_T}, \bar{\eta}\right) = W_{\bar{p_T}} \cdot W_{\bar{\eta}} \tag{4.13}$$

where $W_{\bar{\eta}(\bar{p_T})} = \nu^{D}_{\bar{\eta}(\bar{p_T})} / \nu^{MC}_{\bar{\eta}(\bar{p_T})}$ and ν is the bin content of the normalized distributions.
Such weights are then used as per-event weights on the Monte Carlo events, and the procedure is iterated until these weights stabilize.

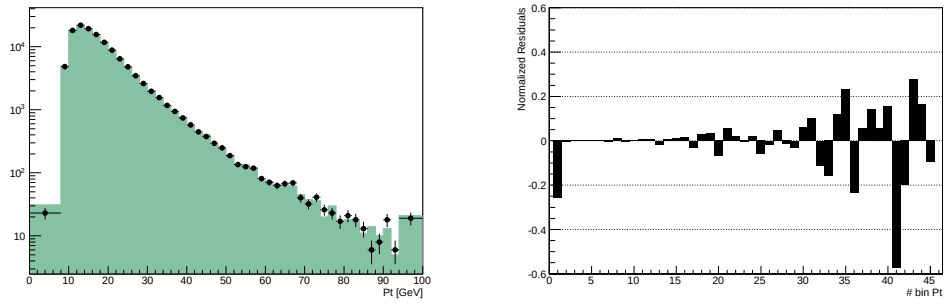
The bin widths of the histograms used is not fixed but determined by the bin content of the data histograms after the sideband-subtraction: bins are dynamically merged until the bin significance, defined as the ratio between the bin content and the bin error, reaches a minimum threshold equal to 2.5.

It is important to underline that this data-driven determination of the weights is subject to the uncertain knowledge of the true values p_T and η of the B meson in data. The size of this effect is driven by the relative size of the resolution in these variables, compared to the width of the bins used in the study. These widths have been studied on Monte Carlo, and the outcome is shown in figures 4.12 and 4.13, respectively for η and p_T: the left plots show the comparison between the Monte Carlo truth distributions (filled green)



(a) η MC distributions (b) Residuals

Figure 4.12: Left: comparison between the truth (filled green) and reconstructed (black point) B^+ η distributions in Monte Carlo samples. Right: bin by bin residuals between the truth (filled green) and reconstructed (black point) B^+ η distributions in Monte Carlo samples.



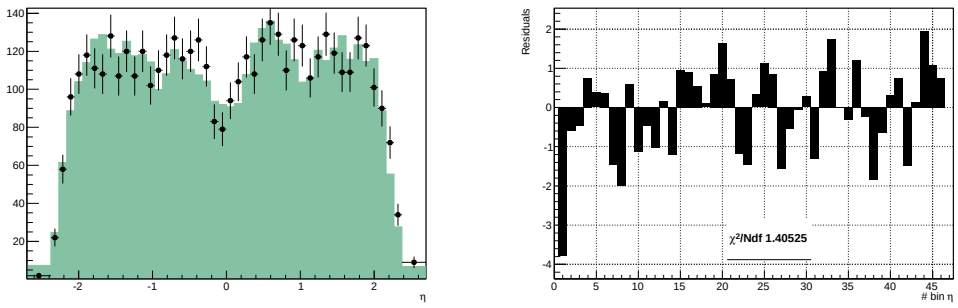
(a) p_T MC distributions (b) Residuals

Figure 4.13: Left: comparison between the truth (filled green) and reconstructed (black point) B^+ p_T distributions in Monte Carlo samples. Right: bin by bin residuals between the truth (filled green) and reconstructed (black point) B^+ p_T distributions in Monte Carlo samples.

and the corresponding reconstructed distributions, while in the right plots the residuals, defined as the bin by bin difference of these two distribution normalized to the bin content of the former distribution are shown. It has been also verified by comparing the Monte Carlo truth and the reconstructed quantity, that within the kinematical regions used, i.e. p_T >6 GeV and $|\eta|$ < 2.5, the bin to bin migration due to the use of measured p_T, η rather than the truth values is within 1/100 of the bin content, therefore negligible with respect to the statistical uncertainties with which the weights themselves are determined.

Convergence test

The re-weighting procedure so far described has been validated with a test using the Monte Carlo sample only; it was split into three not equally populated parts, namely A, B and C. The basic idea of this test is to reproduce the re-weighting procedure as we do in the analysis: the "A" sample distributions were artificially distorted in order to emulate the ones we find in Monte Carlo for the reference channel, the "B" sample was used as a data sample for the reference to employ in the re-weighting, while "C" sample was used as cross reference sample that may emulate a well tuned Monte Carlo.

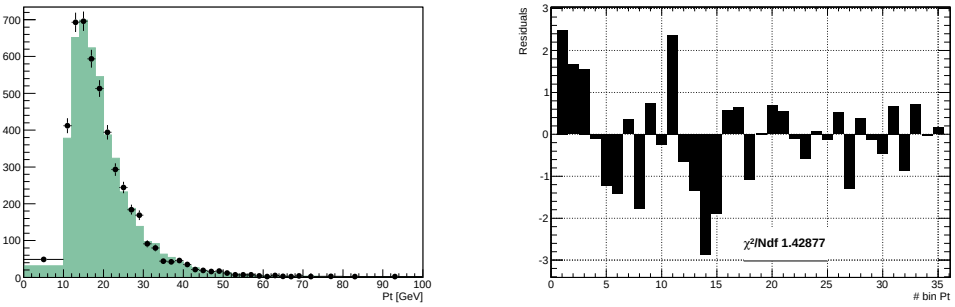


(a) η (b) Residuals

Figure 4.14: Comparison between C (filled green) and B (black points) Monte Carlo η distributions (left) and residuals (right)

.

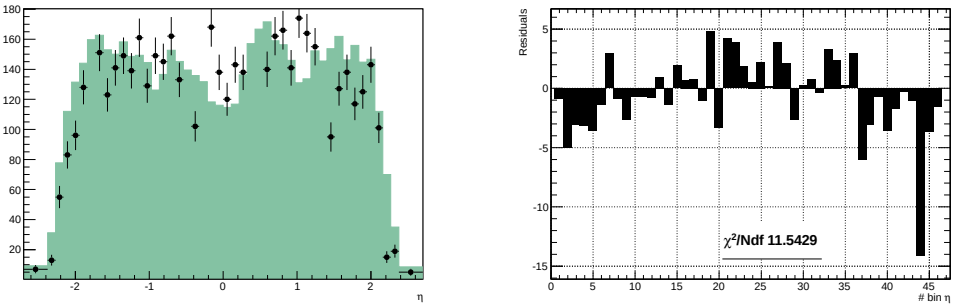
The comparison between the B sample and the C sample as a function of η and p_T are shown, respectively, in figures 4.14 and 4.15: left plots show the distributions of the two samples, where the C sample entries are normalized to the ones in B sample, while the right plots represent the residuals between



(a) p_T (b) η

Figure 4.15: Comparison between C (filled green) and B (black points) Monte Carlo p_T distributions (left) and residuals (right).

them [2]. In both cases the value of the χ^2 test is near to one, confirming that the two distributions are in good agreement and belong to the same dataset. The same comparison between the C and the distorted A Monte



(a) η (b) Residuals

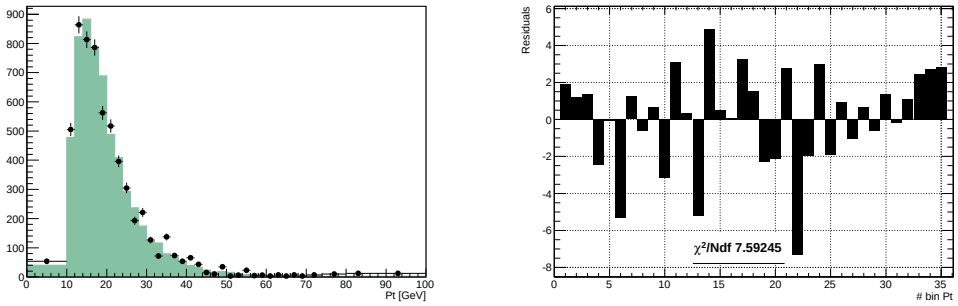
Figure 4.16: Comparison between C (filled green) and A (black points) Monte Carlo η distributions (left) and residuals (right) before the re-weighting.

.

Carlo samples before the re-weighting is shown in figures 4.16 and 4.17; the disagreement between the shapes of the two samples for both η and p_T distributions is evident and confirmed by the high value of the χ^2 test.

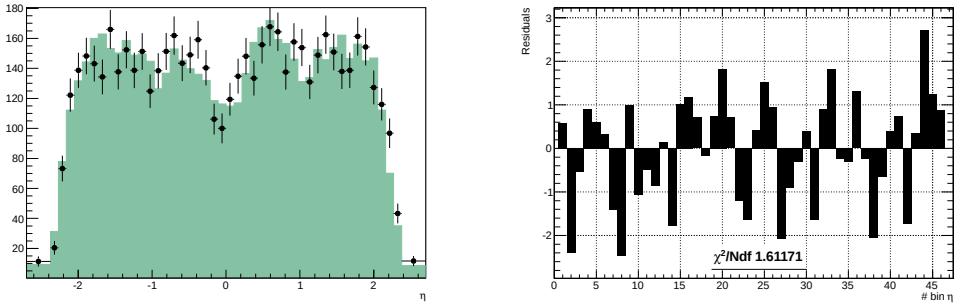
The re-weighting procedure has then been applied: from the comparison of the A and B sample a first set of weights has been computed and used to retune A sample: the output of this first iteration is shown in figures 4.18

[2]To compute residuals and the χ^2 no errors on Monte Carlo C sample distributions were assumed.



(a) p_T (b) η

Figure 4.17: Comparison between C (filled green) and A (black points) Monte Carlo p_T distributions (left) and residuals (right) before the re-weighting.

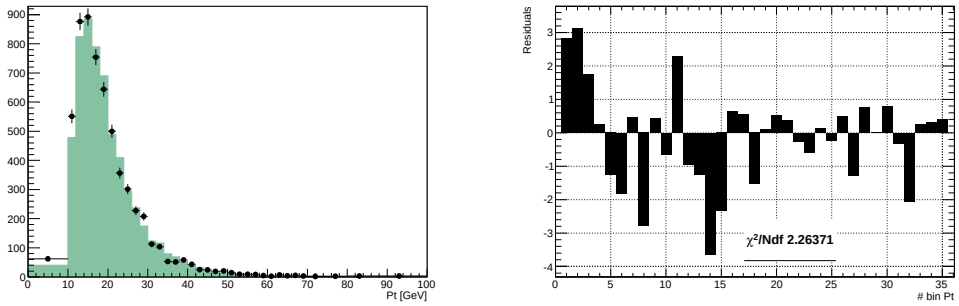


(a) η (b) Residuals

Figure 4.18: Comparison between C (filled green) and A (black points) Monte Carlo η distributions (left) and residuals (right) after the 1^{st} iteration of the re-weighting.

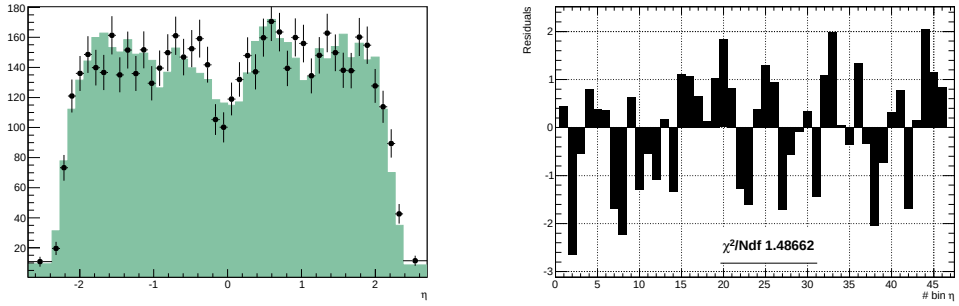
and 4.19 where the η and p_T re-weighted distributions are again compared to the C sample. Already after the first iteration the agreement between the distribution is good, as validated by the χ^2 test result.
A second iteration has been then applied on A sample and the result, reported in figures 4.20 and 4.21, confirms that the method converges; moreover, the χ^2 test values are very similar to to one obtained in the initial comparison between B and C samples.

Another prove of the convergence can be found in figure 4.22. The plots in this figure show, respectively the η and the p_T weights as obtained after two iterations: black points correspond to the weights obtained after the first iteration, thus comparing the distorted distribution and the reference Monte Carlo distributions, while the red points are obtained after the application



(a) p_T (b) η

Figure 4.19: Comparison between C (filled green) and A (black points) Monte Carlo p_T distributions (left) and residuals (right) after the 1^{st} iteration of the re-weighting.



(a) η (b) Residuals

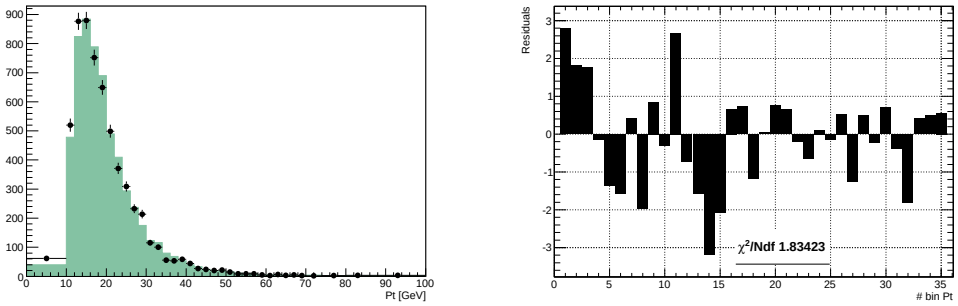
Figure 4.20: Comparison between C (filled green) and A (black points) Monte Carlo η distributions (left) and residuals (right) after the 2^{nd} iteration of the re-weighting.

of the weights to the distorted distributions. The convergence results clear since in the second round the weights stabilize around the value one.

This convergence and the consistency of the result indicates the correctness of assumption that p_T and η are uncorrelated as well as of the binning chosen.

Background subtraction

The $B^{\pm} \to J/\psi K^{\pm}$ data distributions used to re-weight Monte Carlo samples are affected by the contamination of background, mostly composed by non resonant μ^+ μ^- pairs associated to a third track. This kind of contam-



(a) p_T (b) η

Figure 4.21: Comparison between C (filled green) and A (black points) Monte Carlo p_T distributions (left) and residuals (right) after the 2^{nd} iteration of the re-weighting.

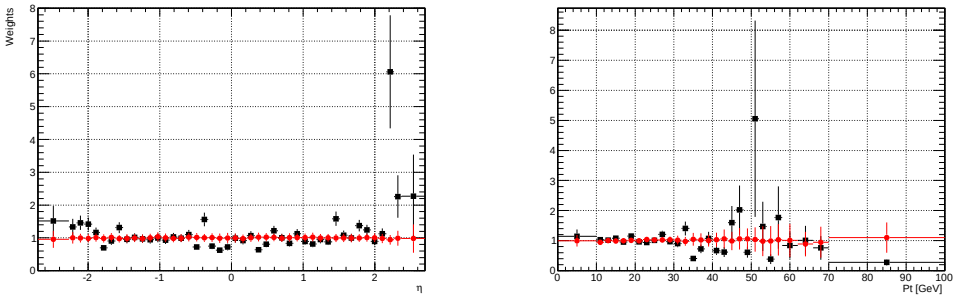
ination can then bias the determination of the weights used for the tuning and if possible has to be removed by subtracting it from the $B^{\pm}$ peak region. To estimate and subtract this kind of background from our samples a data driven method was used which is based on the assumption that non resonant background behavior is the same in both signal and sidebands invariant mass regions, so that sidebands can be exploited to interpolate the number of background events under the $B^{\pm}$ peak.
Thereof, the invariant mass spectra was divided in three regions:

- the signal region, defined in the range [5182.96, 5382.96] MeV,
- the two sidebands, defined in the regions [5082.96, 5182.96] MeV and [5382.96,5482.96] MeV respectively for the left and for the right sideband.

An un-binned maximum likelihood fit is performed on the invariant mass of the $B^{\pm}$ by using the same model described in section 4.2.

The number of signal (N_S) and background events in the peak region (N_B), and the number of background events in both left (N_{LS}) and right (N_{RS}) sideband were obtained from the integral of the fit functions in these regions.

Given the assumption on the background behavior, in order to subtract the background from the signal region, events in the sidebands were assigned a weight $W_{SB} = -N_B/(N_{LS} + N_{RS})$, while events in the signal region were assigned a weight $W_{SB} = 1$. This assignment was leaded by the fact that when an event happens to be in the sidebands it is most probably a background



(a) η weights in two iterations

(b) p_T weights in two iterations

Figure 4.22: η (left) and p_T (right) weights as obtained from two subsequent iterations: black points correspond to first iteration while red points to the second iteration.

event, while no such statement can be done when it falls in the signal region since it can be either signal or background.

Calculation of η and $\mathrm{p_T}$ Weights

Once performed the sideband background subtraction, the weights can be calculated.

As previously shown in 4.3.3, two iterations were needed to have stabilized weights to use in the event re-weighting. The expression of these weights after the first and the second iteration for $\mathrm{p_T}$ reads:

$$W^1_{p_T} = \frac{\sum_\eta D_{\eta,p_T}}{\sum_{\eta,p_T} D_{\eta,p_T}} \frac{\sum_{\eta,p_T} MC_{\eta,p_T}}{\sum_\eta MC_{\eta,p_T}} \tag{4.14}$$

and

$$W^2_{p_T} = \frac{\sum_\eta D_{\eta,p_T}}{\sum_{\eta,p_T} D_{\eta,p_T}} \frac{\sum_{\eta,p_T} MC_{\eta,p_T} W^1_\eta W^1_{p_T}}{\sum_\eta MC_{\eta,p_T} W^1_\eta W^1_{p_T}} \tag{4.15}$$

where in the sum p_T and η are intended as the index of the corresponding bins of the distribution, the normalization factor for the first iteration is

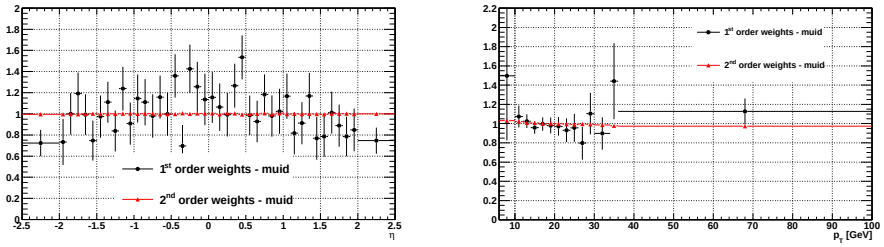
$$\frac{\sum_{\eta,p_T} MC_{\eta,p_T}}{\sum_{\eta,p_T} D_{\eta,p_T}} \tag{4.16}$$

while the normalization factor for the second iteration, taking under consideration the reshape of the Monte Carlo spectra due to the first order correction of the weights, is

$$\frac{\sum_{\eta,p_T} D_{\eta,p_T}}{\sum_\eta MC_{\eta,p_T} W^1_\eta W^1_{p_T}} \,. \tag{4.17}$$

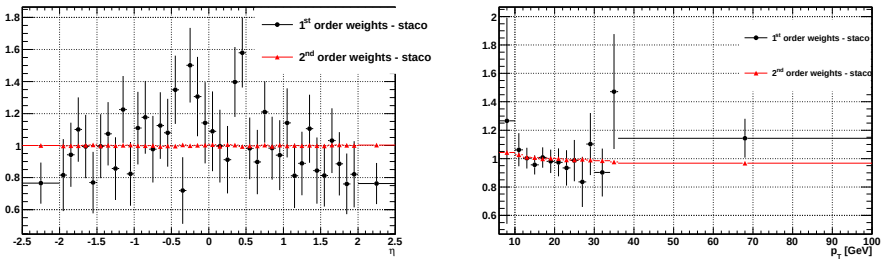
A similar expression holds for the η weights, so the event weight reads:

$$W(p_T, \eta) = W^1_{p_T} \cdot W^2_{p_T} \cdot W^1_{\eta} \cdot W^2_{\eta} \tag{4.18}$$



(a) η weights (b) p_T weights

Figure 4.23: 1^{st} and 2^{nd} order η (a) and p_T (b) weights, muid family. The method converges as indicated by the fact that the second iteration weights are almost equal to one.

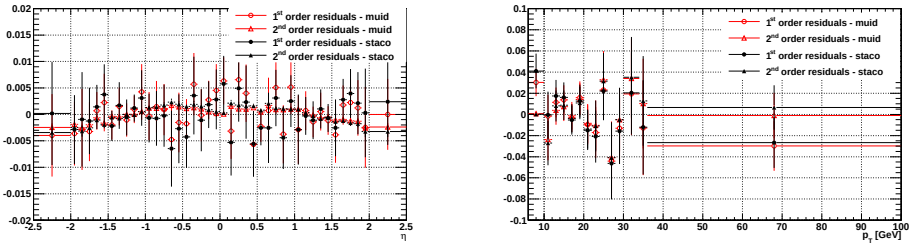


(a) η weights (b) p_T weights

Figure 4.24: 1^{st} and 2^{nd} order η (a) and p_T (b) weights, staco family. The method converges as indicated by the fact that the second iteration weights are almost equal to one.

All possible sources of statistical uncertainties in weights calculation have been taken under consideration, as the uncertainties on the sideband weights, on the Monte Carlo spectra and on the data spectra (including the whole procedure of obtain them).

The errors on the weights have been calculated with a gaussian propagation but the resulting formula is very complicated and also involve directly the MC_{η,p_T} and D_{η,p_T} matrices in the calculation. A software package, which



(a) η weights (b) p_T weights

Figure 4.25: 1^{st} and 2^{nd} order residuals for η and p_T weights for muid and staco families; in both samples the second iteration leads to η and p_T weights converging to one.

has the capability of propagating the errors from the initial spectra up to the calculation of weights, has been implemented on purpose.

From the plots in figures 4.23 and 4.24, showing, respectively, the η (left) and p_T (right) weights after the first and the second iteration for the two muons reconstruction families, the convergence is clear.

Also the comparison of the residuals [3] between data and Monte Carlo re-weighted distributions for the two muon reconstruction families, reported in figure 4.25, confirms that the method converges and gives "universal" correction weights for Monte Carlo samples.

4.4 B_s decay channel selection

4.4.1 Mass resolution families

The muon momentum resolution determines the invariant mass resolution, which is one of the fundamental quantity in this analysis. Muons originating from the B_s meson decay cross different layers of the ATLAS Inner Detector: as a result the muon momentum resolution depends on the pseudo-rapidity. The ATLAS detector is segmented in *Barrel* and *End Cap* but this definition in terms of η depends on the particular detector and the distance from the IP; the segmentation used in this study, schematically reported in figure 4.26, is the following:

- $0 < |\eta| < 1.0$ or "common barrel" (B) (characterized by the fact that

[3] The residuals in this case are defined, for each bin i, as $r_i = \frac{\left(nu_i^D - \nu_i^{MC}\right)}{\nu_i^D}$, where ν_i is the bin content of either data D or re-weighted Monte Carlo MC normalized distributions.

all the three technologies composing the ATLAS Inner Detector belong to the ATLAS barrel);

- $1.0 < |\eta| < 1.5$ or "common transition" (T);
- $1.5 < |\eta| < 2.5$ or "common end-caps" (E).

Using the Monte Carlo signal sample, the di-muon mass resolution has been studied by using this segmentation and three main "resolution" families have been defined as following

1. both muon tracks detected in the common barrel (BB);
2. one track detected in the common barrel, the other one in the common transition or both in the common transition ($BT + TT$);
3. at least one of the two track detected in the common end-caps (E).

The mass resolutions corresponding to this splitting are shown in figure 4.27 as a function of the absolute value of the B_s pseudo-rapidity; the BB sample contains 43.5% of the initial sample and to it corresponds the best mass resolution ($\sim$ 50 MeV). The latter rapidly degrades to $\sim$ 75 MeV for the $BT + TT$ sample (24.8% of the total sample) and to $\sim$ 100 MeV for the E sample.

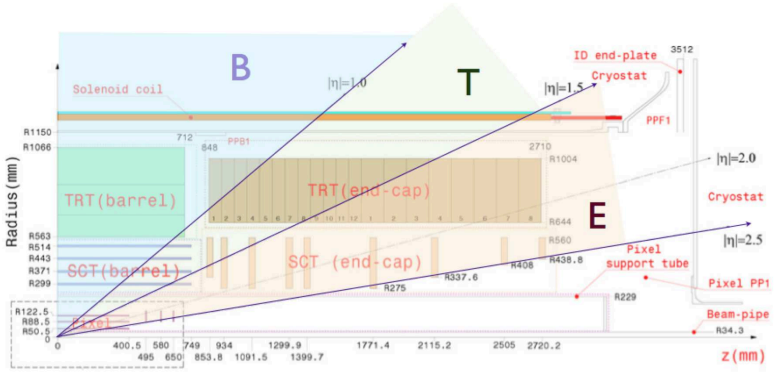
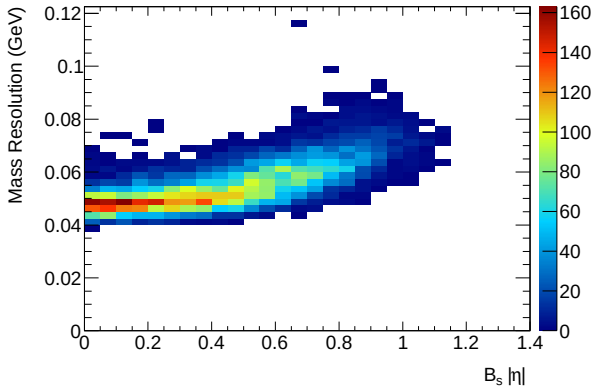


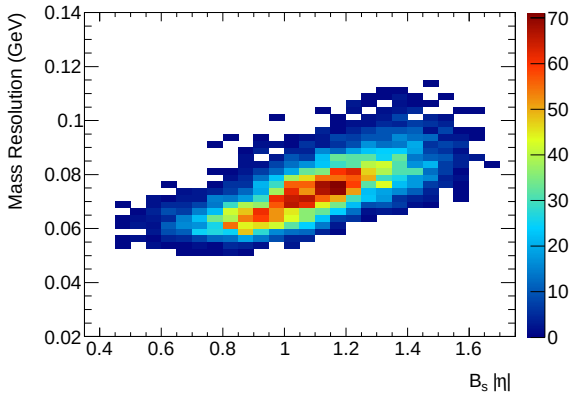
Figure 4.26: Inner detector segmentation as a function of the pseudo-rapidity.

Depending of the available statistics, the splitting of the sample according with the resolution family can be used to achieve better results by using different options as, for example:

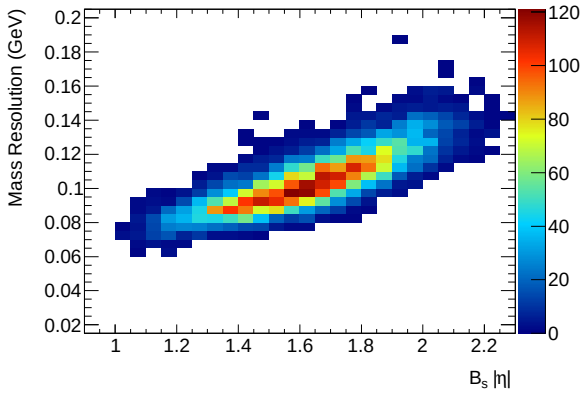
1. use only the BB sub-sample;



(a) BB



(b) BT+TT



(c) E

Figure 4.27: B_s mass resolution as a function of $|\eta|$ of the B_s for the three resolution families.

2. use the full sample and then optimize the mass cut as a part of the analysis;

3. split the sample in three parts, perform separate optimization and finally merge the results.

Although the level of information increases when passing from the first to the third hypothesis, and so the possible gain in terms of signal/background separation, the last solution resulted the most time consuming, so in the optimization the second option has been used.

4.4.2 Signal channel selection

In order to clean-up the initial data sample and also reduce it to a manageable size, both Monte Carlo and data in the sidebands a preliminary event selection has been applied. These "skimmed" samples have been employed then in the optimization procedure.

As for the reference sample, the data in the sidebands come from $\sqrt{7}$ TeV $p-p$ collisions taken during 2010 when the detector was declared to be fully operational and LHC beam stable; the data were selected by lowest non-topological di-muon trigger threshold within the B_s mass window (cfr. sec. 3.2.1) and corresponds to an integrated luminosity of $\sim 40\ pb^{-1}$.

The additional cuts applied on the events are summarized in the following

- **Collision Selection**: to veto cosmic muons, events were required to contain at least one reconstructed primary vertex with at least three associated Inner Detector tracks;

- **Track quality**: tracks used for the B_s meson reconstruction were required to have at least one hit in the Pixel detector, six hits in the SCT an one hit in the TRT tracker;

- **Muon Selection**: events were required to contain at least one couple of opposite charged muons, either formed by two Combined muons or by one Combined and one Segment-Tagged muon (cfr. sec. 3.1.6). The Inner Detector tracks associate with the reconstructed muons should have matched the Track quality requirements; moreover, the highest-p_T muon in the pair should have at least a transverse momentum $p_T > 4$ GeV, while the other at least a transverse momentum $p_T > 2.5$ GeV.

- Bs **Selection**: opposite signed di-muon tracks pass this selection if they are successfully fitted to a common vertex with $\chi^2/ndf < 2$ and if their invariant mass is in the range $[4.763, 5.963]$ GeV. B_s candidates

are required to be reconstructed within the pseudo-rapidity range $|\eta| \leq 2.5$ and to have a transverse momentum $p_T > 6$ GeV.

4.4.3 Signal selection optimization

The signal selection optimization strategy relies on the signal features. As already mentioned before, given its long lifetime the Bs meson travels in the detector for an appreciable distance before decaying. The muons emitted result isolated and the sum of their transverse momenta lies at a small angle with respect to the Bs. The variables used for the discrimination are

1. the B_s meson pseudo-proper time $c\tau$, defined as

$$c\tau = \vec{L}_{xy} \cdot \vec{p}_T / M_{B_s}$$

   where p_T is the B_s candidate reconstructed transverse momentum, M_{B_s} the invariant mass and L_{xy} is the transverse decay length defined as in equation 4.5.

2. the B_s meson pointing angle to the primary vertex in 2D, α, that corresponds to the angle between the transverse momentum $\vec{p}_T$ and the vector pointing from the primary vertex to the B vertex computed in the transverse plane $\vec{X}_{PV}$;

3. the B_s meson isolation, defined as

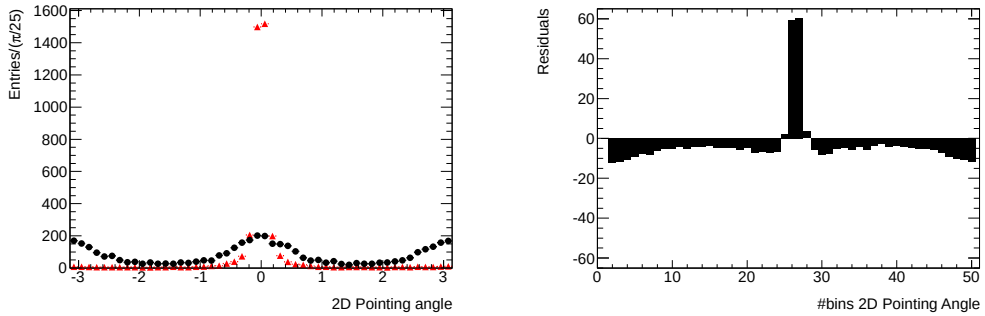
$$I_r = \frac{p_T}{p_T + \sum_{i=0}^{n_{tracks}} p_T^i}$$

   where p_T^i is the transverse momentum carried by the i-th track of the n_{tracks} found in a cone of radius r around the B_s momentum.

Some studies have been performed to evaluate the gain in separation power when switching from 2D to 3D tracking information in the computation of the $c\tau$ and α; the (small) gain achieved did not counterbalance the additional effects introduced in the MC reproducibility of such variables.

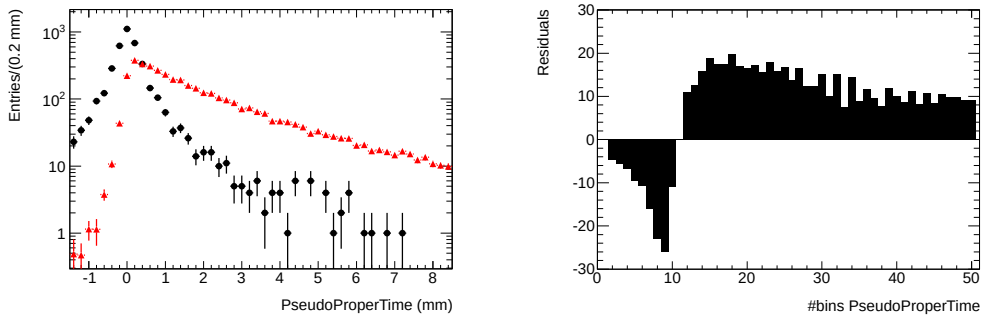
Monte Carlo samples have been used to model signal distributions of the selection variables, although the reliability of such model has already been called into question and the re-weighting techniques described in section 4.3 had to be used.

Non-resonant background is the predominant source of background so, in order to model it data in the signal mass sidebands have been used, where the sidebands range in the intervals [4766.33,5066.33] MeV (left sideband)



(a) 2D Pointing angle (b) Residuals

Figure 4.28: 2D Pointing angle distribution (left) and residuals(right): Monte Carlo signal (red triangles) is compared to non resonant background (black dots) from data sidebands. MC distribution is normalized to data entries.

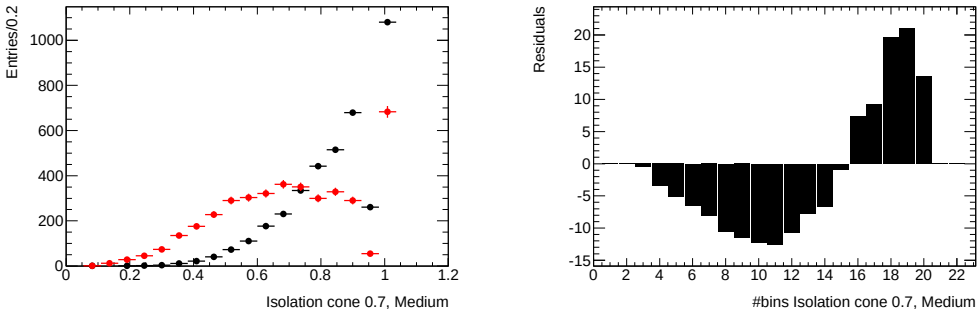


(a) Pseudo proper time (b) Residuals

Figure 4.29: Pseudo proper time distribution (left) and residuals(right): MC signal red triangles) is compared to non resonant background (black dots) from data sidebands. MC distribution is normalized to data entries.

and [5666.33,5966.33] MeV (right sideband). The width of the two sidebands has been chosen symmetric and in such a way to avoid mass ranges belonging to other possible di-muon resonances (ψ', $\Upsilon(1S)$).

A qualitative feeling for the behavior and the separation power of the variables so far can be obtained by looking at figures 4.28, 4.29 and 4.30, where signal MC (red triangles) and data (black points) sidebands, together with the residuals between the two distributions are shown.

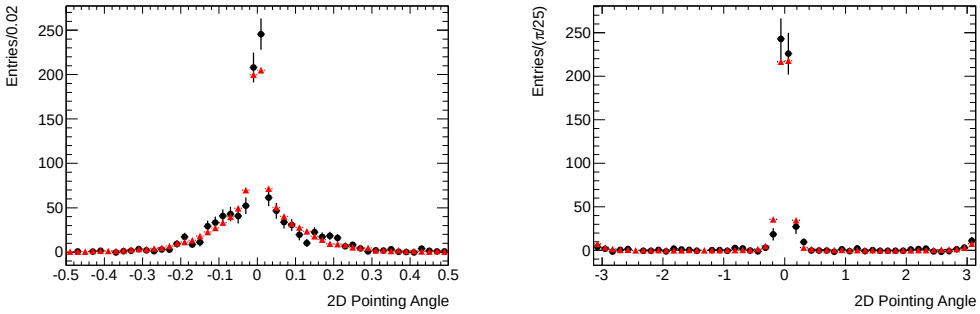


(a) Isolation (b) Residuals

Figure 4.30: Isolation (cone 0.7) distribution (left) and residuals(right): Monte Carlo signal (red triangles) is compared to non resonant background (black dots) from data sidebands. MC distribution is normalizd to data entries.

4.4.4 Monte Carlo validation

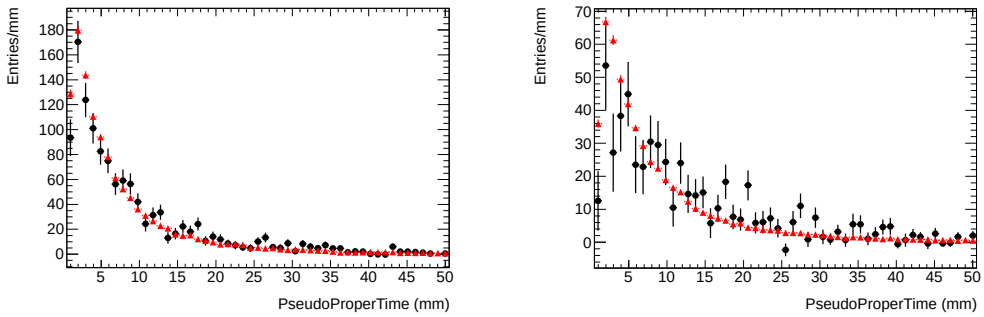
The analysis strategy foresees the use of Monte Carlo samples to both evaluate the efficiencies that enter in the SES and optimize the signal selection. Although the re-tuning procedure has been fully validated by using the relevant B mesons kinematic variables (i.e. the transverse momentum and the pseudo-rapidity), further checks have been done for the signal selection variables.



(a) B^+ 2D Pointing angle (b) J/ψ 2D Pointing angle

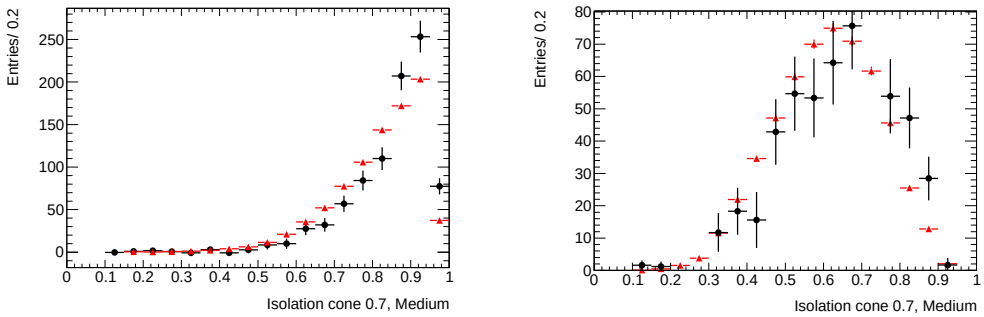
Figure 4.31: 2D Pointing angle comparison between data (black) and re-weighted Monte Carlo samples for B^+ (a) and J/ψ (b) candidates

Particularly interesting for this comparison could have been the $B_s \to J/\psi(\mu^+\mu^-)\phi(K^+K^+)$ channel which has a branching ratio only 1/2 lower than the $B^\pm \to J/\psi K^\pm$ one (see reference [69]). This decay channel, in fact, has a B_s in the initial state and then the muons kinematics is most likely the



(a) B^+ Pseudo proper time (b) J/ψ Pseudo proper time

Figure 4.32: Pseudo proper time comparison between data (black) and re-weighted Monte Carlo samples for B^+ (a) and J/ψ (b) candidates



(a) B^+ isolation (b) J/ψ isolation

Figure 4.33: Isolation (cone 0.7) comparison between data (black) and re-weighted Monte Carlo samples for B^+ (a) and J/ψ (b) candidates

signal and moreover also the f_u/f_s ratio would in this case cancels out.

This possibility has been evaluated but the limited statistics available in 2010 (only $\sim$ 436 candidates were found in ATLAS data [82]) made it impossible to use and thereof the more abundant $B^\pm \to J/\psi K^\pm$ sample has been employed.

In particular, both the 3-prong $\mu\mu K$ (corresponding to the $B^\pm$) and the 2-prong $\mu\mu$ (corresponding to the J/ψ) vertices belonging to the same sample have been used on this purpose; while the former allows to test the general B meson kinematic variables as, for instance, the isolation, the latter allows to check quantities more related to vertex multiplicity (e.g. the decay-length resolution).

Some examples of the comparison between sideband subtracted data and Monte Carlo re-weighted distributions are reported in figures 4.31, 4.32,

4.33, where left plots show, respectively, the 2D pointing angle, pseudo-proper time and isolation for the B^+ while right plots show the corresponding variables for J/ψ. Indeed, particular care should be used for the isolation as

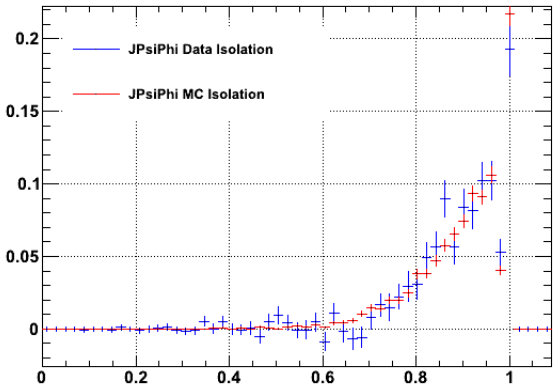


Figure 4.34: Isolation (cone 0.7) comparison between data and re-weighted Monte Carlo samples for $B_s^0 \to J/\psi\phi$ candidates in 2011 data.

its modeling in the MC is not the most reliable one; this is due to its intrinsic dependence on the modeling of the b quark hadronization into a B meson, which makes it dependent on the B specie. Therefore in the $B^{\pm}$ selection no isolation cut has been applied and then the systematic uncertainty on the isolation cut efficiency on the B_s candidates should be evaluated separately. Indeed, studies performed on 2011 data shows no discrepancies between data an Monte Carlo samples in the case of $B_s^0 \to J/\psi\phi$ for what concern this variable [4] (cfr. fig. 4.34) and then the systematic uncertainty is neglected.

4.4.5 The Cut and count approach

The baseline strategy for the $B_s \to \mu^+\mu^-$ Upper Limit extraction by using 2010 data is the so-called "Cut & Count" approach in which the extraction of the number of observed event is done by merely applying some cuts on the kinematic variable and count the number of event survived. The method has been chosen for its simplicity and robustness, allowing to give a first estimation of the limit and also of the systematics to take in account. It has been also used as reference for the more complex technique developed for the analysis on 2011 data that, on the other side, are less immediate and prone to bias.

[4] The χ^2 test performed on data ans MC distribution gives $\chi^2/ndf = 1.07$ with a probability of 36%.

In order to improve the limit, an optimization of the cut to apply to the most sensible variable has been performed. The three selection variables previously described (α, $c\tau$, $I_{0.7}$, ΔM), together with the invariant mass window (ΔM) around the B_s mass ($M_{B_s} = (5.3663 \pm 0.0006)$ GeV [69]) have been chosen and their ranges have been divided into 20 equal slices. The events are selected if they falls within a specified range defined by the cut: for the selected variables the range are defined by $|\alpha| < \alpha_{cut}$, $|c\tau| > c\tau_{cut}$, $|\Delta M| < \Delta M_{cut}$, $Iso > Iso_{cut}$.

For each sub-range, the medium point has been used as a possible cut value. The resulting 4×20 grid allows for a contemporarily scan of the selection variables, thus disregarding the correlations among them. By scanning this grid the signal efficiency (ϵ_N) and the number of background events (N_{back}) passing the cuts defined by each grid point has been evaluated, together with the Figure of Merit (FoM) proposed by Punzi [84]:

$$\mathcal{P} = \frac{\epsilon_S}{\frac{a}{2} + \sqrt{N_{bkg}}} \tag{4.19}$$

In expression 4.19, a is a parameter related to the significance of the sensitivity which for $a = 2$ corresponds to 95% Confidence Level.

The optimal value for the cuts chosen corresponds to the maximum value obtained for the Punzi estimator in the scan. The scan has been refined around the optimal points with a finer steps in the variables sub-ranges (see table 4.2) defined by the point where the Punzi estimator dropped to 50% of its maximum (the values used have been rounded to obtain a meaningful binning).

This estimator has been proved to provide the best upper limit when used for the optimization of cutting variables. In fact, it is strictly related to the un-knowledge of the signal branching ratio, as in expression 4.19 the signal efficiency is used rather than the number of selected signal events; this fact, together with the luminosity dependence in the $\sqrt{N_{back}}$ term, prevent optimization algorithms from reducing the background to zero and creating a undesirably low signal yield.

By using a toy Monte Carlo, assuming the same integrated luminosity of 2010 data, the Punzi Figure of Merit has been compared to the widely used Figure of Merit

$$\mathcal{E} = \frac{S}{\sqrt{S+B}}$$

where S and B are, respectively, the signal and the background yield surviving the cuts.

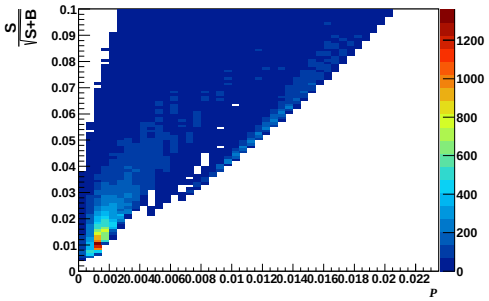


Figure 4.35: Comparison between $\frac{S}{\sqrt{S+B}}$ and Punzi FoM.

The result, shown in figure 4.35, is that the two Figure of Merits are strongly correlated and optimization is pretty much the same.

Rolke Upper Limit (see chapter 5) versus $\mathcal{P}$ has also been checked. Figure 4.36 shows the Upper Limit behavior as a function of $\mathcal{P}$: the best Upper Limit value is correctly obtained around the maximum values of the Punzi's Figure of Merit. The Punzi FoM is then the best choice for the analysis providing a better Upper limit when compared to the others and also enhancing the signal yield in the optimization.

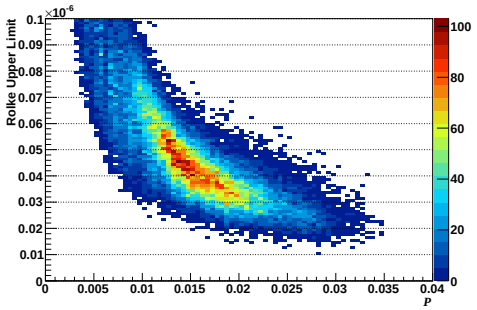


Figure 4.36: Upper Limit á la Rolke as a function of the Punzi FoM, $\mathcal{P}$

α	$c\tau$ [mm]	$I_{0.7}$	ΔM GeV
[0.03-0.83]	[0.2125-0.4295]	[0.42-0.82]	[0.238-0.818]

Table 4.1: Ranges used for the fine scan of the selection variables.

The optimization procedure has been applied to the samples used in the analysis. The scan has been performed by using both STACO and MUID

	α	$c\tau$ [mm]	$I_{0.7}$	ΔM GeV	$\mathcal{P}$	$\epsilon_{\mu\mu}\alpha_{\mu\mu} \times \epsilon_N$	N_{back}
STACO	0.085	0.3	0.71	0.56	0.023	0.046	1
MUID	0.085	0.3	0.71	0.56	0.021	0.042	1

Table 4.2: Optimal values obtained from the scan.

muons reconstruction families. The variables ranges used are reported in table 4.1, while the values obtained for the optimal cuts and the estimator are reported in table 4.2.

4.4.6 Optimization procedure validation

The signal selection optimization procedure so far described relies on a unique data sample, i.e. the odd-numbered events in the signal sidebands. This is due to the fact that the events in the sidebands have also to be used for the estimation of the N_{bgk} of background events and then, to avoid possible bias, these events are splitted in two samples.

The optimization procedure has to be fully validated for the sake of the analysis result, but as by product also provides a valid cross reference for other possible optimization approaches, like multivariate analysis studies.

In order to fully validate this procedure, some effects that may introduce biases in the Upper Limit estimation have been evaluated, like

- the background interpolation bias in optimization;
- the stability of the $\mathcal{P}$ optimal point with respect to statistical fluctuation of the background;
- the evaluation of biases, efficiency and background yield with luminosity.

The technique used to evaluate these effects is the so-called "bootstrap" procedure: each event i, $i = 1, ..., n$ belonging to the M measures set is used as a "seed" to generate an arbitrary number of event sets B_j, $j = 1, ..., m$ via Poisson extraction. The so-obtained event samples are statistically independent samples of the original M sample [5].

[5]The parent distribution of the B_j sets is M itself, not the one which M represents.

Background interpolation bias in optimization and evolution with luminosity

The bootstrap technique has been used to estimate which bias is introduced when using the entire data sample in performing both the optimization of the selection variables and the background interpolation.

Even numbered data events in sidebands events have been used as seeds for this bootstrap, and $N = 1000$ new samples, each containing the same number of events found in the sidebands, have been generated.

On each J sample, $J = 1, ..., N$, the optimization procedure has been applied: the cuts for the optimal value of selection variables have been used to interpolate the number of background passing the cuts N_{back} for both the J (N^J_{back}) and the $J + 1$ (N^{J+1}_{back}) samples.

Plots in figure 4.37 (a) show the distribution for N^J_{back}, $\Delta N = N^{J+1}_{back} - N^J_{back}$ and for the efficiency so obtained. In particular, the ΔN distribution is not centered at zero and this indicates that indeed a bias is introduced.

In order to check the bias evolution with luminosity, this test has been repeated by progressively increasing the bootstrap samples size. The result is shown in figure 4.37 (b) as a function of the number of background events: although with the increasing of the luminosity the bias is quite negligeable, the uncertainty on it remains constant, so for this analysis we decided to keep splitting the sample for the two separate tasks awaiting for new options [6].

4.5 Acceptance and efficiencies evaluation

One of the most relevant quantities in the determination of the SES and the limits extraction is the ratio

$$R = \frac{\epsilon_{B^\pm}\alpha_{B^\pm}}{\epsilon_{\mu\mu}\alpha_{\mu\mu}} \times \frac{1}{\epsilon_N} \tag{4.20}$$

The efficiencies and the acceptances have been determined from the MC samples described in section 4.3.1, after applying the re-weighting procedure described in section 4.3. The products $\varepsilon_x\alpha_x$ have been calculated as a single number; in fact, in this context, each product has to be interpreted as the ratio of candidates passing the full set of selection cuts for the corresponding signal, to the total number of generator-level events produced within a certain p_T, η range chosen to be the same for $B^\pm$ and B_s events, in such a way not to introduce kinematic biases in the selected events. In this way,

[6] Monte Carlo samples could be used for the cut optimization and the full statistics for background interpolation, but was not available for 2010 data analysis.

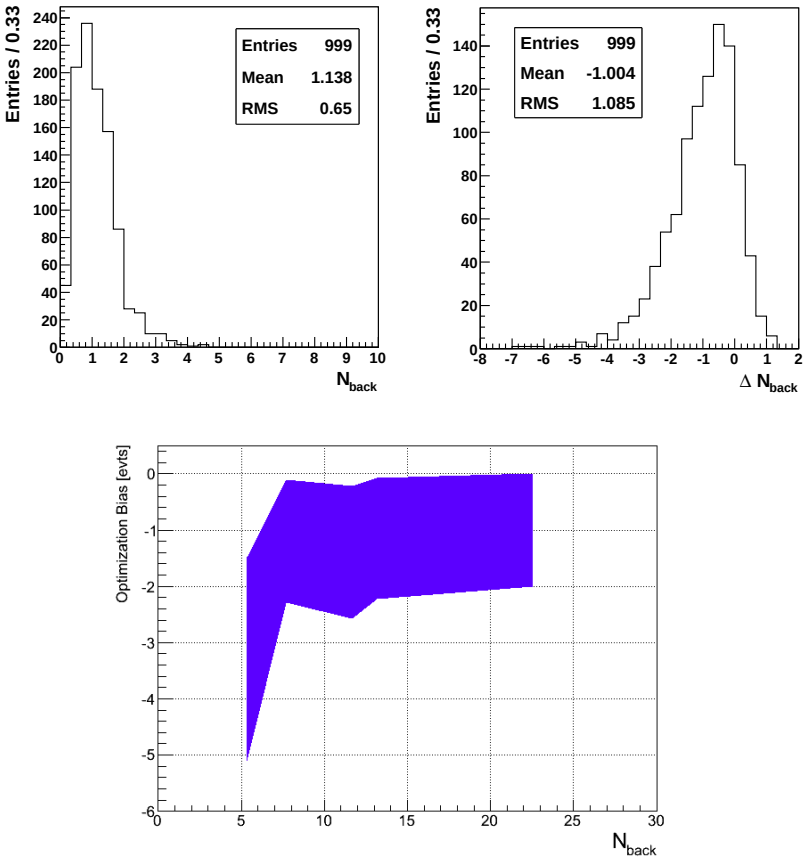


Figure 4.37: (a) Distribution of N^J_{back} (left) and ΔN as obtained by Toy Monte Carlo. (b) Bias evolution as a function of the number of background.

also the uncertainties coming from the re-weighting are more naturally taken into account and they can be propagated to the product straightforwardly. It is important to note that the two product $\epsilon_{B^\pm}\alpha_{B^\pm}$ and $\epsilon_{\mu\mu}\alpha_{\mu\mu}\epsilon_N$ use the same weights and this introduces a correlation in the errors. Once accounted properly, the uncertainty on the ratio is ± 1.5% for both STACO and MUID. Moreover, on this ratio the systematic uncertainty due to the use of B^+ sample to re-tune the B_s Monte Carlo has been evaluated for the 2011 analysis within the B-Physics group, and the outcome of the study is that this effect may range between 3% and 6.5%. Since this second uncertainty is not negligeable with respect to the former, we decided to conservatively take it into account as a 6.5% effect. The result obtained for the ratio R is 0.937± 1.4% (stat) ± 6.5% (syst)% for the STACO family, and to 0.947±1.5%(stat) ± 6.5% (syst) for MUID.

4.6 Single Event Sensitivity (SES) evaluation

STACO		
$N_{B^\pm}$	R	SES/10^{-8}
2754±2%(stat.)±0.8%(syst.)	0.930±1.5%(stat)±6.5%(syst)	7.4±10.6%
MUID		
$N_{B^\pm}$	R	SES/10^{-8}
2834±1.8% (stat.)±0.4%(syst.)	0.947±1.5%(stat)±6.5%(syst)	7.3±10.6%

Table 4.3: Single Event Sensitivity ingredients.

All ingredients so far evaluated, reported in table 4.3 are used to build the SES with 40 pb^{-1} of $p-p$ ATLAS collision data, for both STACO and MUID families. The normalization with respect to the $B^\pm \to J/\psi K^\pm$ decay channel and the re-weighting of the Monte Carlo sample allowed to get rid of most of the systematic uncertainties, the main remaining being the one on the fragmentation functions and on the $B^\pm \to J/\psi K^\pm$ branching ratio, whose product is equal to 2.2×10^{-4}. The uncertainty on the latter quantities, estimated to be $\sim 8.4\%$, reflects directly into an uncertainty on the branching ratio limit.

Thus, given the SM branching ratio $Br(B_s \to \mu^+\mu^-) = (3.2\pm0.2) \times 10^{-9}$, from equations 4.3 and 4.4, we would expect to see at this integrated luminosity a number of signal candidates equal to:

$$N_{\mu\mu}^{STACO} = \frac{Br(B_s \to \mu^+\mu^-)}{SES} = 0.043 \pm 0.005 \tag{4.21}$$

in the case of STACO and

$$N_{\mu\mu}^{STACO} = \frac{Br(B_s \to \mu^+\mu^-)}{SES} = 0.044 \pm 0.005 \tag{4.22}$$

in the case of MUID.

A systematic effect which can be introduced in the SES evaluation when using the $B^+ \to J/\psi K^+$ Monte Carlo sample to evaluate the $\varepsilon_x \alpha_x$ product, since both positive and negative charged $B^\pm$ candidates in data to evaluate the yield.

Charged Kaons interactions with matter, in fact, depend on their electric charge. For example, negative charged kaons can be involved in processes like $K^- N \to Y\pi$, where $Y = \Lambda, \Sigma$...., and N is a proton or neutron, while the corresponding charge conjugate process would only proceed in the presence of

anti-protons or anti-neutrons. This reflects in a lower detection efficiencies for the negative kaons since they can interact with the detector material, resulting in a shortened (or missing) track, that propagates to the $B^{\pm} \to J/\psi K^{\pm}$ efficiency evaluation.

The difference between the detection efficiencies ϵ_K can be expressed in terms of the following asymmetry

$$A_K = \frac{\epsilon_K^+ - \epsilon_K^-}{\epsilon_K^+ + \epsilon_K^-}$$

Studies performed within the B-Physics Rare decay group revealed that this effect is relatively small ($A_K \sim 1\%$), and corresponds to a $\sim 0.5\%$ reduction of the B^+ efficiency; thus it can be accounted in the SES as a systematic uncertainties.

More in details, the kaon charge asymmetry is measured by selecting in ATLAS data $D^{*\pm}$ candidates reconstructed in the $D^0(\mu^{\pm}K^{\mp})\pi^{\pm}$ decay channels.

Assuming that the $D^{*\pm}$ production do not depend on charge, the kaon asymmetry will be evaluated as the asymmetry between the number of positively and negatively charged kaons reconstructed in these chains

$$A_K = \frac{N_K^+ - N_K^-}{N_K^+ + N_K^-}$$

$D^{*\pm}$ candidates are searched by looking at the mass difference ΔM = M($\mu K\pi$) - M(μ K): the small difference between the $D^{*\pm}$ and the D^0 mass ($\sim$ 145MeV) gives a signal peak with relatively small combinatorial background, mainly composed of

1. direct $\mu\pi\pi$ background from $D^0(\mu^{\pm}\pi^{\mp})\pi^{\pm}$ decay;

2. indirect $\mu K\pi$ background from $D^0(\mu^{\pm}K^{*\mp}(K^{\mp}\pi^0))\pi^{\pm}$ decay;

3. indirect $\mu\pi\pi$ background from $D^0(\mu^{\pm}K^{*\mp}(K^0\pi^-))\pi^{\pm}$ decay;

The $\mu\pi\pi$ 3-prong vertex background components, differently from signal, does not present charge asymmetry in the reconstruction efficiency. This corresponds to a dilution of the asymmetry when the combinatorial background is subtracted from data, so the "raw" asymmetry has been corrected in order to account for this effect; the correction factor has been estimated from Monte Carlo as the fraction of reconstructed $\mu K\pi$ vertices $f_K = \frac{N_{\mu K\pi}}{N_{\mu K\pi}+N_{\mu\pi\pi}} = (91.3 \pm 0.1)$ %.

Moreover, since the kaon asymmetry is expected to be p_T dependent, the p_T spectrum of the kaon has been corrected by using the difference between signal Monte Carlo for $D^{*\pm} \to D^0(\mu^\pm K^\mp)\pi^\pm$ and $B^+ \to J/\psi K^+$ samples.

The "raw" asymmetry has been estimated to be

$$A_K^{raw} = -0.0096 \pm 0.0051$$

which, corrected for the dilution, becomes

$$A_K = -0.0105 \pm 0.0056\,.$$

This effect is then very small and can be safely neglected.

4.7 Background evaluation

4.7.1 Non resonant background

As already stated, the main source of background is given by (prompt) Drell-Yan pairs ($pp \to \mu^+\mu^-$) and non-prompt heavy flavor semileptonic decays ($pp \to b\bar{b} \to \mu^+\mu^- X$): this kind of background is evaluated by using the even-numbered events in the sidebands, and then taking into account the fact that half of the statistic has effectively been used.

The number of observed events passing the selection cuts in the sidebands is 6 six for both STACO and MUID families. A simple linear interpolation in the signal region would give an expected number of background events in the signal region but the uncertainty on it that would derive from a fit is too high.

Thus, for the Upper Limit calculation the number of observed events in the sidebands together with the ratio τ between the width of the sidebands region and the signal region will be used as input instead of the estimated number of background events in the signal region.

4.7.2 Irreducible resonant background

In addition to non-resonant background, the irreducible background which originates from neutral B meson two-body decays with charged hadrons in final state ($B_x \to hh'$) has been investigated.

The list of these decay modes with the relative branching ratios Br are reported in table 4.4. Although the branching ratio are 2-3 order of magnitude greater than the one expected for the signal channel, the probability of mis-identify a final state hadron as muon either due to punch-through or

Decay mode	Br/10^{-6}	Ref
$B_s \to K^+K^-$	(35±7)	[69]
$B_s \to \pi^+K^-$	(5.0±1.0)	[69]
$B_s \to \pi^+\pi^-$	(0.57±0.15(stat.)±0.10(syst.))	[71]
$B_d \to K^+K^-$	(0.23±0.10(stat.)±0.10(syst.))	[71]
$B_d \to K^+\pi^-$	(19.4±0.6)	[69]
$B_d \to \pi^+\pi^-$	(5.13±0.24)	[69]

Table 4.4: $B \to hh'$ decays branching ratios.

	Fake rate/10^{-3}		
Hadron	Combined	Segment-Tagged	Combined or Segment-Tagged
K	(3.94±0.16)	(1.66±0.10)	(5.60±0.19)
K^+	(4.29±0.24)	(2.14±0.17)	(6.44±0.29)
K^-	(3.58±0.22)	(1.16±0.12)	(4.74±0.25)
π	(1.93±0.11)	(0.78±0.07)	(2.71±0.13)
π^+	(1.88±0.15)	(0.89±0.10)	(2.77±0.18)
π^-	(1.98±0.15)	(0.67±0.09)	(2.65±0.18)

Table 4.5: Muon fake rate for various hadron flavors in the various muon types, Monte Carlo only.

decay in flight (*muon fake rate*) has been estimated to be $\sim 10^{-3}$. Since these fake muons carry the most part of the initial hadron momentum, their invariant mass may fall into the signal mass region. In the following the estimation of the yield coming from this background with respect to the signal yield by using only STACO muons is described, as in the case of MUID the final result would be the same within the uncertainties.

The muon fake rates from both Combined, Segment-Tagged and Combined or Segment-Tagged cases estimated by Monte Carlo are reported in table 4.5. The difference between K^+ and K^- fake rates derives from the kaon charge asymmetry in the interactions with the detector described before.

The results for Combined fake rates have been compared to the ones obtained by using data-driven techniques; this comparison is reported in table 4.6 and in the case of kaons good agreement between data and Monte Carlo is shown.

For each $B_x \to hh'$ decay mode the yield has been evaluated, as for the case of the signal, with a reference normalization approach.

Hadron	fake rate/10^{-3} (MC)	fake rate/10^{-3} (DATA)
K	(5.60±0.19)	(8.7±1.3)
K^+	(6.44±0.29)	
K^-	(4.74±0.25)	
π	(2.71±0.13)	(0.6±0.4)
π^+	(2.77±0.18)	
π^-	(2.65±0.18)	

Table 4.6: Muon fake rate for various hadron flavors, Combined or Segment-Tagged only.

In particular, an expression similar to the one on equation 4.3 can be used for each of these modes, e.g.:

$$Br(B_x \to hh') = \frac{N_{B_x}}{N_{B^\pm}} \times \frac{\epsilon_{B^\pm}\alpha_{B^\pm}}{\epsilon_{B_x}\alpha_{B_x}} \times \frac{1}{\epsilon_N}\frac{f_u}{f_x} \times Br(B^\pm \to J/\psi K^\pm) \times Br(J/\psi \to \mu^+\mu^-) \quad (4.23)$$

and by dividing equations 4.3 and 4.23 one obtains

$$\frac{Br(B_s \to \mu^+\mu^-)}{Br(B_x \to hh')} = \frac{N_{B_s}}{N_{B_x}}\frac{\epsilon_{B_x}\alpha_{B_x}}{\epsilon_{B_s}\alpha_{B_s}}\frac{f_x}{f_s} \quad (4.24)$$

where $x = s, d$ and $\epsilon_{B_x}\alpha_{B_x}$ accounts also for the h (h') fake rate $\epsilon^{fake}_{h(h')}$. By assuming that the acceptance, trigger and reconstruction efficiencies are the same for both $B_s \to \mu^+\mu^-$ and $B_x \to hh'$ modes, equation 4.24 reduces to

$$\frac{Br(B_s \to \mu^+\mu^-)}{Br(B_x \to hh')} = \frac{N_{B_s}}{N_{B_x}}\epsilon^{fake}_h \times \epsilon^{fake}_{h'}\frac{f_x}{f_s} \quad (4.25)$$

$$N_{B_x} = \frac{1}{SES}Br(B_x \to hh') \times \epsilon^{fake}_h \times \epsilon^{fake}_{h'}\frac{f_x}{f_s} \quad (4.26)$$

For each decay mode, the ingredients entering in equation 4.26 are shown in table 4.7. The total yield of peaking expected background with respect to the $B_s \to \mu^+\mu^-$ has so been estimated to be 3.7×10^{-3}, with a statistical uncertainties of $\pm\ 0.3\times 10^{-3}$ evaluated by taking into account the errors on the fake rates. The systematic uncertainties entering are

- errors on fake rates evaluated with Monte Carlo samples;
- errors on branching ratios.

Decay mode	$\mathrm{Br}/10^{-6}$	$\epsilon_h^{fake} \times \epsilon_{h'}^{fake}/10^{-6}$	$\frac{f_x}{f_s}$
$B_s/\bar{B}_s \to K^+K^-$	35	30.5	0.113
$B_s \to \pi^+K^-$	2.5	13.1	0.113
$\bar{B}_s \to K^+\pi^-$	2.5	17.1	0.113
$B_s/\bar{B}_s \to \pi^+\pi^-$	3.8	7.3	0.113
$B_d/\bar{B}_d \to K^+K^-$	0.23	30.5	0.401
$B_d \to \pi^+K^-$	9.7	13.1	0.401
$\bar{B}_d \to K^+\pi^-$	9.7	17.1	0.401
$B_d/\bar{B}_d \to \pi^+\pi^-$	5.13	7.3	0.401

Table 4.7: Ingredients and resonant background yield evaluation.

The first systematic uncertainty, evaluated as difference with respect to the data driven fake rates reported in table 4.6, is equal to $\pm 0.04 \times 10^{-3}$.
The second, instead, entails the uncertainty of $\pm\ 3 \times 10^{-3}$ to the events yield. These two combined together give a systematic uncertainties of $\pm 3 \times 10^{-3}$ and the final number of resonant background with respect to the $B_s \to \mu^+\mu^-$ results $N_{back}^{res} = \left(5.2 \pm 0.3_{(stat.)} \pm 3.0_{(syst.)}\right) \times 10^{-3}$; thus, this background contribution will be neglected in the Upper Limit extraction that will be threated in the next chapter.

Chapter 5

The Upper Limit on the $B_s \to \mu^+\mu^-$ branching ratio

In this chapter the Upper Limit extraction on the $B_s \to \mu^+ \mu^-$ branching ratio will be presented. The achieved result will be compared to the latest results from the other experiments to have a hint on where and how the limit can improve with a larger statistics. The 2011 analysis, which is currently in the final stage and makes use of $\sim$2.4 fb^{-1} of integrated luminosity, will be briefly described.

5.1 Introduction

In the study of a rare decay particular care should be used in quoting the result of the measurement, as generally the model for the signal is either unknown or knows with high uncertainty. Once a model is assumed, the problem of setting a limit on the number of the observed signal events falls back into the general problem of giving a statement concerning the agreement or the disagreement between the observed data and the predicted model.

When making inferences about some unknown parameters of a model, for which a functional form is known, a statistical function of the dataset (which takes into account the chosen model) is used to determine them within a certain range. The estimation of the parameters comes together with a Confidence Level (CL), i.e. the probability that by repeating the experiment the outcome will fall in the same range; statistical procedures are used also to evaluate whether data support or not a particular hypothesis, and also in this case a CL is quoted.

The way in which these procedures are built is dictated by the approach one wants two adopt; the two main approaches are the Frequentist and the

Bayesian [69], although there are other methods, like the likelihood L, which do not correspond to either of these two.

Indeed, the likelihood approach is one of the most used tool to make inferences about a data sample. This technique is generalized into more complex methods when evaluating at the same time more than one hypothesis or to include the so-called "nuisance parameters", i.e. all the quantities that may intervene in the observation (e.g. detector resolution, reconstruction efficiency,..) and need to be included in the estimation although do not give more information concerning the model itself.

The outcome of an experiment does not always lead to a measurement, especially in the rare processes cases where the probability to observe them it's very low. In these cases, rather than just saying that nothing was found, a statement concerning the range of the excluded values for the signal is given as it can help in confirming or refuting a certain theory.

The Upper Limit setting is complicated by the presence of uncertainties on the signal and background models. The typical scenario is a counting experiment where the number n of observed events, expected to be distributed according to the Poisson law, should be compared to the expected events in the background (b) or signal plus background ($s + b$) hypothesis. In order to distinguish among the two cases a statistic test q is used and the two hypothesis can be expressed, respectively, as $f(q|b)$ and $f(q|s + b)$.
Depending on the observed data, the statistic provides a value, q_{obs}, which is used to determine the confidence intervals. The output of the statistic q_{obs} can be used to also evaluate the probability of an eventual disagreement with the prediction to be due only to a statistical fluctuation through the p-value. The latter is defined as

$$p_b = P(q \leq q_{obs}|b) = \int_{-\infty}^{q_{obs}} f(q|b)dq \tag{5.1}$$

in the b hypothesis and as

$$p_{s+b} = P(q \geq q_{obs}|s + b) = \int_{q_{obs}}^{\infty} f(q|s + b)dq \tag{5.2}$$

in the case of signal plus background hypothesis. Note that the $p-$value provides only a measurement of the disagreement with the expected model, but does not account for a possible mistake in the model itself.

There are several methods to set Upper Limits, and usually on the same sets of data they may provide different results; the most common approaches to set Upper Limits are the following:

1. Feldman Cousin (FC) [83];

2. Rolke [85];

3. CL_s [86].

The first is a fully frequentistic method and it's built in order to guarantee the coverage [1]; the ordering rule in building the confidence intervals is the likelihood ratio:

$$Q = \frac{L(x;\mu)}{L(x;\mu_{best})} \tag{5.3}$$

where x corresponds to the observed parameter, μ to the true parameter and μ_{best} to the value of μ that maximizes L.

The Feldman Cousin approach should be used in case of no (or negligible) uncertainties since it is not capable of treating uncertainties in nuisance parameters; therefore, it is not used in this analysis to extract the Upper Limit because we want to take into account the uncertainties in the SES and in the background estimation.

Rolke method is based on a "profile likelihood" approach, where the profile likelihood is a likelihood which accounts for the nuisance parameters θ, i.e. $L = L(x;\mu,\theta)$ and is evaluated at the value of the nuisance parameters $\hat{\theta}$ which maximize it for each μ. The confidence levels are built according to the ratio of the profile likelihood:

$$\lambda(\mu) = \frac{L(x;\mu\hat{\theta})}{L(x;\mu_{best},\theta_{best})} \tag{5.4}$$

where μ_{best} and θ_{best} are the values which maximize L. The resulting intervals for the parameters of interest are not guaranteed to have the exact coverage probability for all values of the nuisance parameters, but in cases of practical interest the approximation is found to be very good. The signal in this approach is assumed to be distributed according to the Poisson law, while different assumptions can be made on the distribution of the nuisance parameters.

The CL_s method is defined starting from the p-value definitions and extended the so-called "CL_{s+b}" method where the statistical test of the $s+b$ hypothesis is based on p_{s+b}: the signal model is excluded at a confidence level of 1-α = 95% if one finds:

$$p_{s+b} < \alpha. \tag{5.5}$$

In the CL_s procedure, instead, the signal is excluded if:

$$CL_s = \frac{p_{s+b}}{1-p_b} < \alpha. \tag{5.6}$$

[1] The coverage is a property of a method that specifies how often the confidence interval contains the true value of interest.

Upper Limits on the mean value of a Poisson or Gaussian distributed measurement obtained by using CL_s are the same as in the Bayesian Upper Limits. This feature render the CL_s a suitable tool to easily combine results from different experiments.

In this analysis the Rolke method will be used to quote the Upper Limit, while the CL_s will be used to counter-check the obtained result.

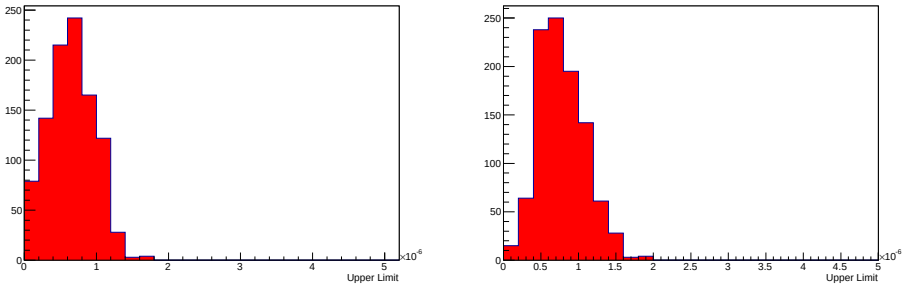
5.2 Upper Limit extraction

5.2.1 The Rolke Upper Limit

The Rolke method has been implemented within the ROOT [67] framework, and uses as inputs for the Upper Limit determination

- the inverse of the Single Event Sensitivity with its uncertainty;
- the number of observed events in the signal mass region;
- the number of background events in the sidebands;
- the ratio τ between the total mass width of the sidebands and the signal mass window.

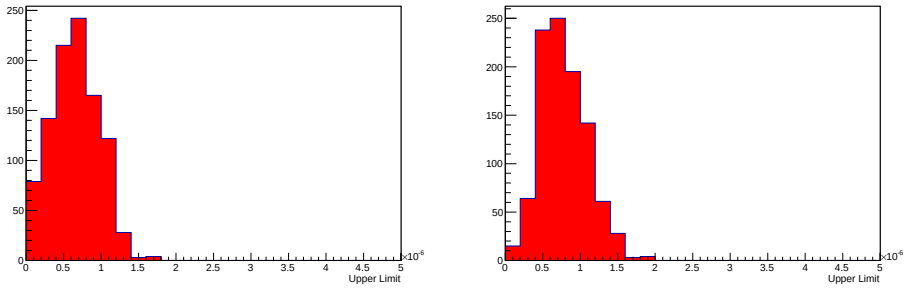
In particular, the inverse of the SES corresponds to the efficiency for the signal; we set a gaussian model for the uncertainty on this quantity while for the background we assumed a poissonian model. In particular, the last two



(a) Upper Limit (90% CL) (b) Upper Limit (95% CL)

Figure 5.1: Toy experiment results on Upper Limits estimated by using STACO muon family @ 90% (a) and 95% (b) CL.

inputs serve to estimate the number of the events in the sidebands with the relative uncertainties and fold them in the Upper Limit extraction.



(a) Upper Limit (90% CL) (b) Upper Limit (95% CL)

Figure 5.2: Toy experiment results on Upper Limits estimated by using MUID muon family @ 90% (a) and 95% (b) CL.

	Mean (90%CL)	Median (90%CL)
STACO	7.24×10^{-7}	6.57×10^{-7}
MUID	7.14×10^{-7}	6.48×10^{-7}
	Mean (95%CL)	Median (95%CL)
STACO	8.75×10^{-7}	7.91×10^{-7}
MUID	8.63×10^{-7}	7.80×10^{-7}

Table 5.1: Mean and median expected Upper Limits obtained by using toy experiments for both STACO and MUID muon families.

Before the un-blinding, 1000 toy experiments have been performed to evaluate the Upper Limit expectations at 90% and 95% CL by using for the SES the value reported in table 4.3, the number of background events in the sidebands (6 in both STACO and MUID case) and the ratio $\tau = 0.54$. All these ingredients, but the number of observed event (obtained via Poisson extraction) are the same used in determine the observed Upper Limit; the outcome of these experiments, shown in figures 5.1 and 5.2, respectively for STACO and MUID families, is reported in table 5.2.1.

The values in the table confirm that the Upper Limit extraction depends on the sensitivity of the experiment: in particular, since the number of observed background events in the sidebands (and the corresponding uncertainties) is the same for both STACO and MUID (cfr. sec. 4.7), the slightly smaller value of the SES in the MUID case (cfr. tab. 4.3) determines a better value for the Upper Limit.

The Upper Limit depends also on the accuracy on the SES; besides the systematic uncertainties coming from the product of the fragmentation func-

tions ratio and of the $B^\pm \to J/\psi K^\pm$ branching ratio, most of the uncertainty on the SES are statistical, as we choose to re-weight Monte Carlo samples according to data. Thus, although the expected Upper Limit results higher than the one imposed by the other experiments, it will improve as the integrated luminosity increases.

5.2.2 Final result

Once the analysis has been approved and the selection cuts freezed, the latters have been applied on the un-blinded data sample for both MUID and STACO families. In the signal region, corresponding to a mass window of $\pm$ 0.28 GeV around the B_s mass we observe, for both muon families, nine events: the di-muon invariant mass in a range that includes the sidebands is shown in figure 5.3.

	Upper Limit at 90%CL	Upper Limit at 95%CL
STACO	$4.75\times\ 10^{-7}$	$6.00\times\ 10^{-7}$
MUID	$4.69\times\ 10^{-7}$	$5.91\times\ 10^{-7}$

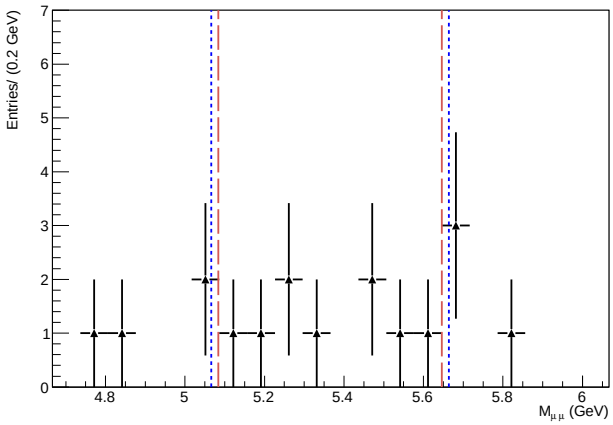
Table 5.2: Observed Upper Limits obtained for STACO and MUID families.

This result has so been used to compute the Upper Limit on the number of observed signal events determined by using the Rolke procedure: the results are reported in table 5.2. The CL_s method, again implemented within the

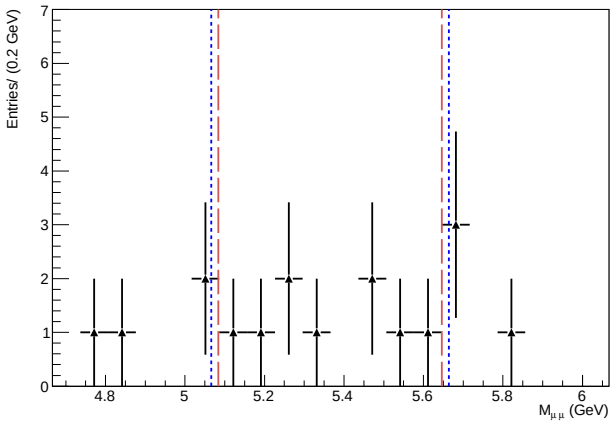
	Observed Limit	Expected Upper Limit (Median)
STACO	$7.01\times\ 10^{-7}$	$7.65\times\ 10^{-7}$
MUID	$7.01\times\ 10^{-7}$	$7.89\times\ 10^{-7}$

Table 5.3: Observed and expected Upper Limits at 95% CL obtained by using CL_s method for STACO and MUID families.

ROOT framework, has been also used to counter-check these results. The Upper Limit has been obtained by using the same input as in the Rolke case. The input likelihood is built by assuming a poissonian PDF for both signal plus background $(s + b)$ and background only (b) hypothesis, and a scan of the p-values by varying the expected branching ratio for the signal has been used to compute the limit. The results, reported in table 5.3 are shown also in figure 5.4. Although slightly less constraining with respect to the ones



(a) STACO Invariant mass



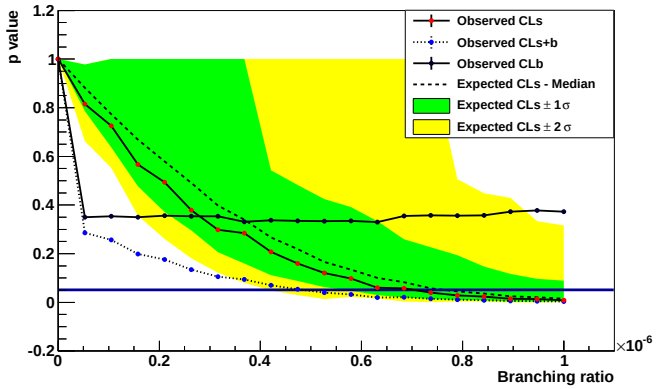
(b) MUID Invariant mass

Figure 5.3: Di-muon Invariant mass after the data unblinding for STACO (a) and MUID (b) muon families; the dotted points separarate the signal region from the sidebands.

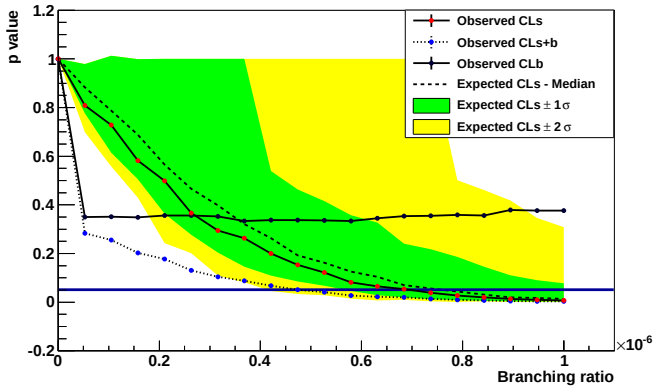
obtained with the Rolke method, the CL_s Upper Limits have the advantage to be directly compared to the one obtained with the other experiments.

5.3 Comparison with other experimental results

The result on the Upper Limit obtained in this analysis can be compared to the latest results quoted by the other LHC experiments, LHCb [49] and CMS [50], and to the one quoted by CDF [48] experiment at TEVATRON



(a) STACO CL_s Upper Limit



(b) MUID CL_s Upper Limit

Figure 5.4: Upper Limits at 95% CL obtained by using the CL_s method for STACO (a) and MUID (b) muon families. The Upper Limit on the Standard Model branching ratio is set when the observed CL_s curve goes below a $p-$value of 0.05.

that, in the mean time, provided also the branching ratio measurement $Br(B_s \to \mu^+\mu^-) = 1.8^{+1.1}_{-0.9} \times 10^{-8}$ (although the latter has not been confirmed neither by LHC nor from the other TEVATRON experiment, D0).

All these results, reported in table 5.4 together with the integrated luminosity used to extract them, are obtained by using the CL_s method. The best Upper limit correspond to the combination of the LHCb an CMS results [51], which is $Br(B_s \to \mu^+\mu^-) < 1.08 \times 10^{-8}$ at 95% CL, and refutes the CDF measurement of the branching ratio.

The main difference between the Upper Limit obtained in this analysis

Experiment	Integrated Luminosity (fb^{-1})	Upper Limit at 95% CL
LHCb	0.337	1.5×10^{-8}
CMS	1.14	1.9×10^{-8}
CDF	7	4×10^{-8}

Table 5.4: Upper Limits at 95% CL obtained on the $B_s \to \mu^+\mu^-$ branching ratios by different experiments.

and the others just mentioned is the integrated luminosity used to extract the result: in fact, the used statistic is $\sim$ 175 times higher in the case of CDF, $\sim$25 in the case of CMS and $\sim$ 8.4 in the case of LHCb. Moreover, even with the same integrated luminosity the comparison with LHCb would be unfair since this experiment has been designed and optimized to study B Physics.

Indeed, in order to provide a competitive limit, the ATLAS Collaboration decided to optimize the analysis by using 2011 data rather then presenting the result obtained with the baseline analysis.

Disregarding the final result, a qualitative comparison can be made on the analysis strategies adopted by the other experiments.

As in this analysis, the "blind search" and the evaluation of the branching ratio by using a normalization sample, in order to cancel in the ratio the systematic uncertainties, are used. In particular, both CMS and CDF experiments use as reference channel $B^{\pm} \to J/\psi K^{\pm}$ decay channel, while thanks to its features LHCb can use in addition also the $B_s^0 \to J/\psi\phi$ and the $B^0 \to K^+\pi^-$ decays.

The main difference between the analysis described in this thesis and the CMS one is that, in the second case, the sample has been split in two parts according to the mass resolution that varies as function of the pseudorapidity; then the two results are used to build a likelihood with the SM branching ratio as a common constraint. As in this analysis, CMS adopts a cut and count approach and optimizes the selection cuts in order to obtain the best Upper Limit; in this optimization the signal efficiency is evaluated by using Monte Carlo and validated by using as control sample the $B_s^0 \to J/\psi\phi$ decay channel, while the non resonant background has been modeled by using sidebands in data. From the paper is not clear if the same (full) sideband sample has been used to both perform the cut optimization and background interpolation thus possibly introducing a bias in the background determination (cfr. sec. 4.4.6). Although CMS reaches may achieve higher resolution than ATLAS in resolving the di-muon invariant mass, we do expect to obtain

a similar result on the Upper Limit by using a similar integrated luminosity.

LHCb analysis is quite more sophisticated: after the selection, in both signal and normalization samples, to each event is given a probability to be signal or background like; this probability is obtained by using a two-dimensional likelihood built by using two independent variables: the invariant mass and the response of a multivariate classifier, the Boosted Decision Tree (BDT) from the TMVA [87] package.

Finally, the same approach of building a multi-dimensional likelihood in bins of mass and classifier is used by CDF, this time the classifier used is a Neural Network (NN).

These comparisons allowed to individuate two main improvements for the ATLAS analysis strategy that have been adopted for the 2011: the use of multivariate classifier to improve the signal/background separation, and the use of a multi-dimensional likelihood built by using both classifier output and mass resolution families spaces; more detail will be given in the next section.

5.4 The 2011 data analysis

The 2011 analysis makes use of $\sim$2.4 fb^{-1} of $p-p$ collision data collected up to July 2011; up to that period, in fact, the trigger used to select events for both reference and signal channel have been kept un-prescaled. Many features of the analysis strategy result unchanged with respect to 2010 baseline analysis (cfr. sec. 4.1), but some improvements have been introduced.

Both the blind-analysis search and the normalization of the branching ratio with respect to the $B^\pm \to J/\psi K^\pm$ decay channel are adopted as in the baseline analysis.

The fitting technique in the case of the reference channel has been improved by including in the background model also the partially reconstructed decays.

In order to discriminate signal (modeled with Monte Carlo) from background events (taken from half of the sidebands in data), a Boost Decision Tree has been used as estimator and its output is cross-checked to the one from the multi-dimensional cut scan; several variables are added to train the BDT estimator, although when ranked by discriminating power the ones used in this analysis are the best (cfr. sec. 4.4.3).
Among them, the Isolation variable, whose definition is based on the number of tracks coming from the same B candidate vertex, is the one that is mainly affected by the increased pile-up due to the increasing of the luminosity in 2011; in order to cope with this, an optimized definition of the isolation, based on a classifier that allows to discriminate tracks coming from

the primary vertex, has been adopted.

The efficiency and acceptance are still evaluated by using Monte Carlo; this time the re-weighting procedure according to data is used to evaluate the systematic uncertainties on them as a good agreement between data and Monte Carlo is observed; moreover, the increased integrated luminosity allows to use also the $B_s^0 \to J/\psi\phi$ as control sample to validate the Monte Carlo.

All these improvements, together with the increased integrated luminosity, allowed to obtain a single event sensitivity of $\sim 10^{-9}$.

Finally, in the Upper Limit extraction a multi-dimensional likelihood functions in bins of mass and BDT output is used, and the Upper Limit will be evaluated with the CL_s method.

Most of the analysis is done and the unblinding is scheduled to happen in few weeks from now; then we cannot show here any results even if a consistent amount of time has been devoted to this analysis.

Conclusions

This thesis work has been focused on the study of the rare decay $B_s \to \mu^+\mu^-$ with the ATLAS detector at LHC. This is a very rare decay for which the Standard Model predict a branching ratio of $Br(B_s \to \mu^+\mu^-)=(3.2\pm0.2)\times10^{-9}$, which can significantly be enhanced by New Physics processes contributions.

$p-p$ collision data, collected at a center of mass energy of $\sqrt{s} = 7$ TeV during 2010 and corresponding to an integrated luminosity of ~ 40 pb^{-1}, have been used to quote an Upper Limit on the branching ratio.

The search for a rare decay is generally complicated because of different reasons, the main being the lack of knowledge on the theoretical model, that un-avoidably entails the measurement with a systematic uncertainty, and the difficulty in detection itself. This means that the analysis strategy should be particularly accurate, needs to take into account all the possible source of systematic uncertainties and should avoid any possible bias.

This lead to choose a "blind analysis" strategy, meaning that the signal region in data corresponding to an invariant mass of $\pm$ 300 MeV around the B_s mass has been excluded from the data analysis. The signal is modeled then relying on Monte Carlo samples but at cost of introducing some systematic uncertainties in the measurement.

Studies in the determination of the Upper Limit have been performed only after all the other uncertainties and systematic error evaluation were finished.

In order to reduce these systematic uncertainties, the $B_s \to \mu^+\mu^-$ branching ratio has been evaluated as a ratio with respect the branching ratio of another well measured B meson decay, featuring almost the same kinematics as the signal: the "reference" channel. In this case, as reference the $B^\pm \to J/\psi K^\pm$ decay channel was chosen in virtue of its relative abundance in data, and branching ratio reads

$$Br(B_s \to \mu^+\mu^-) = \frac{N_{\mu\mu}}{N_{B^+}} \times \frac{\mathcal{L}_{B^+}}{\mathcal{L}_{\mu\mu}} \times \frac{\epsilon_{B^+}\alpha_{B^+}}{\epsilon_{\mu\mu}\alpha_{\mu\mu}} \times \frac{f_u}{f_s} \times Br(B^\pm \to J/\psi K^\pm) = SES \times N_{\mu\mu}.$$

The Single Event Sensitivity (SES) represents the sensitivity of the experiment to the measurement being independent on the number of event (i.e.

the integrated luminosity), containing all the elements to determine the BR.

The luminosity ($\mathcal{L}$) factors in the ratio cancel out since the luminosity used for both channel is the same, taking also into account the trigger selection efficiencies. Thus, the branching ratio elements that needed to be evaluated are

1. the number of expected signal events $N_{\mu\mu}$;
2. the number of reference channel candidates in data, found to be $N_{B^+} \sim 2800$ candidates by using a maximum likelihood fitting technique;
3. the ratio $\epsilon_{B^+}\alpha_{B^+}/\epsilon_{\mu\mu}\alpha_{\mu\mu}$, that accounts for the B^+ and B_s relative efficiencies (ϵ) and acceptances (α), estimated from Monte Carlo;
4. the ratio of the fragmentation function f_u/f_s ;
5. the $B^\pm \to J/\psi K^\pm$ branching ratio.

For the last two elements in the list, the values used are quoted, respectively, by the LHCb [70] and Particle Data Group [69]. These quantities, together with the uncertainties on the number of signal events, constitute the main sources of systematic uncertainties on the SES as the most of the remaining, coming out from the non correct modeling of the B meson kinematics in data or of the detector response, were converted into statistical uncertainties by re-tuning both B_s and B^+ Monte Carlo samples according to data.

The signal selection, based on a "cut and count" approach on the variables that showed the best discrimination power, was optimized in order to achieve the best Upper Limit. The estimator used in the optimization was the Punzi's [84] Figure of Merit, which has the advantage to be independent on the (still unmeasured) signal cross section.
In this optimization the re-tuned Monte Carlo was used as reference for the signal, while the main sources of background, coming from combinatorial Drell-Yan processes, were modeled according to data distribution in the sidebands of the signal region. Indeed, following the prescription to avoid possible bias, the sample was split in two parts: one was used to optimize the cuts, while the other was used to estimate the number of background events in the signal region.

The elements so evaluated were used as input for the Upper Limit extraction by using the Rolke method, which resulted

$$Br(B_s \to \mu^+\mu^-) < 5.91 \times 10^{-7} @ 95\% CL.$$

The CL_s method, commonly used by LHC experiments, provided, instead, an Upper Limit of

$$Br(B_s \to \mu^+\mu^-) < 7.01 \times 10^{-7} @ 95\% CL.$$

Although not competitive with the latest experimental results, these limits are valuable since obtained by using a low integrated luminosity and a rather simple approach.

In fact, it has been used as baseline and cross check also for 2011 data analysis, which is still on going and in which have been introduced some improvements, consisting mainly in the use of a multivariate estimator to optimize the signal/background separation and in the limit procedure extraction. These improvements, together with the much higher integrated luminosity (~ 2.4 fb^{-1}), will improve the limit, making the ATLAS result very competitive and determinant in approaching Standard Model prediction for the branching ratio.

Bibliography

[1] L. Evans, *New J. Phys.* ***9*** *335* (2007).

[2] The ATLAS Collaboration, *JINST* ***3*** *S08003* (2008).

[3] The CMS Collaboration, R. Adolphi *et al.*,*JINST* ***3*** *S08004* (2008).

[4] The LHCb Collaboration, A. Alves *et al.*,*JINST* ***3*** *S08005* (2008).

[5] The ALICE Collaboration, K. Aamodt *et al.*,*JINST* ***3*** *S08002* (2008).

[6] http://www.stelab.nagoya-u.ac.jp/LHCf/LHCf/index.html.

[7] http://totem.web.cern.ch/Totem/.

[8] http://press.web.cern.ch/press/PressReleases/Releases2008/PR09.08E.html.

[9] *LHC Design Report CERN-2004-03* (2004).

[10] LHC commissioning, http://lhc-commissioning.web.cern.ch/lhc-commissioning/.

[11] A. Moraes *et al.*, *Eur. Phys. J. C* ***50****, 435-466* (2007).

[12] A. Djouadi, *Phys.Rept.* ***457****:1-216,2008*, hep-ph/0503172.

[13] LEP Working Group for Higgs boson searches, *Physics Letters B* ***565****:61-75* (2003).

[14] CDF and D0 Collaborations, *CDF Note 10606* and *D0 Note 6226* (2011).

[15] The ATLAS Collaboration, *ATLAS-CONF-2011-135*(2011).

[16] The CMS Collaboration, *CMS-PAS-HIG-11-011* (2011).

[17] The ATLAS Collaboration, *ATLAS-CONF-2011-157* (2011).

[18] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TopPublicResults.

[19] https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOP.

[20] The ATLAS Collaboration, *ATLAS-CONF-2011-120* (2011).

[21] The CMS Collaboration, *CMS-PAS-TOP-11-008* (2011).

[22] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/BPhysPublicResults .

[23] https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsBPH .

[24] http://lhcb.web.cern.ch/lhcb/Physics-Results/LHCb-Physics-Results.html .

[25] The LHCb COllaboration, *LHCb-CONF-2011-056* (2011).

[26] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, *Physics Letters B* ***429*** *263*(1998).

[27] T. Kaluza, *Z. Phys., 895-906* (1926).

[28] I. Antoniadis, *Physics with Large Extra Dimensions*, *CERN's Yellow Report*(2001).

[29] The ATLAS Collaboration, arXiv:1109.6606v1.

[30] The CMS Collaboration, https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsSUS.

[31] The ATLAS Collaboration, *ATLAS-CONF-2011-097*, arXiv:1108.1316v1.

[32] The CMS Collaboration, https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsEXO11039.

[33] S. L Glashow, *Nucl. Phys.* **22**:579 (1961).

[34] A. Salam in *Elementary Particle Theory: Groups and Analycity*, pag.367 (ed. Svartholm); Almquist and Wiksell(Stocholm) (1968).

[35] S. Weinberg, *Phys. Rev. Lett.***19**: 1264 (1967).

[36] C. S. Wu *et al. Phys. Rev.* ***105****: 1413-1415* (1957).

[37] P. W. Higgs, *Phys. Lett.* **12**:132 (1964).

[38] P. W. Higgs, *Phys. Rev. Lett.* **13**:508 (1964).

[39] P. W. Higgs, *Phys. Rev. Lett.* **145**:1156 (1966).

[40] E. Fermi, *Z Phys.***88**: 161 (1934); E. Fermi, *Nuovo Cimento* **11** (1934).

[41] J. Goldstone *Nuovo Cimento* **19**: 15 (1961).

[42] S.L Glashow; J. Iliopoulos, L. Maiani, *Phys. Rev. D* **2**: 1285 (1970).

[43] N. Cabibbo, Phys. Rev. Lett. **10**:531 (1963).

[44] A. Buras, *Weak Hamiltonian, CP Violation and Rare Decays*, arXiv:hep-ph/9806471v1 (1998).

[45] T. W. Yeh, H. Li, Phys. Rev. **D 56**:1615 (1997).

[46] K.G. Wilson, Phys. Rev. **179**: 1499 (1969).

[47] C. Bobeth, T. Ewerth, F. Kruger, J. Urban, *Analysis of neutral Higgs-boson contributions to the decays* $\bar{B}_s \to l^+l^-$ *and* $\bar{B} \to Kl^+l^-$, *Phys. Rev. D* ***64****:074014* (2001).

[48] CDF Collaboration, *Search for* $B_s^0 \to \mu^+\mu^-$ *and* $B_d \to \mu^+\mu^-$ *with CDF II* arXiv:1107.2304 .

[49] The LHCb Collaboration, *Search for the rare decays* $B_{(s)}^0 \to \mu^+\mu^-$ *with 300* pb^{-1} *at LHCb*, LHCb-CONF-2011-037 (2011).

[50] The CMS collaboration, *Search for* $B_s^0 \to \mu^+\mu^-$ *and* $B \to \mu^+\mu^-$ *decays in pp collisions at* $\sqrt{7}$ *TeV*, CERN-PH-EP-2011-120, submitted to Phys. Rev. Lett (2011).

[51] CMS and LHCb Collaborations, *Search for the rare decays* $B_s^0 \to \mu^+\mu^-$ *at the LHC with CMS and LHCb experiments*, http://cdsweb.cern.ch/record/1374913 .

[52] S. P. Martin, *A Supersymmetry Primer*, [hep-ph/9709356v6] (2011).

[53] P. H. Chankowski, L. Slawianowska, $B_{d,s} \to \mu^+\mu^-$ *in the MSSM*, hep-ph/0008046.

[54] A.G. Akeroyd, F. Mahmoudi, D. Martinez Santos, *The decay* $B_s \to \mu^+\mu^-$*: update SUSY constraints and prospects* arXiv:1108.3018v1 (2011).

[55] C. Beskidt, W. de Boer, D.I. Kazakov, F. Ratnikov, E. Ziebarth, V. Zhukov, *Constraints from the decay $B_s \to \mu^+\mu^-$ and LHC limits on Supersymmetry*, arXiv:1109.6775v1.

[56] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/PixelPublicResults.

[57] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/InDetTrackingPerformanceApprovedPlots#Tracking_Performance_ATLAS_CONF.

[58] The ATLAS Collaboration, *Eur.Phys.J.C* ***70****:787-821* (2010).

[59] The ATLAS Collaboration,*Eur.Phys.J.C* ***70****:723-753*(2010).

[60] The ATLAS Collaboration, *CERN-PH-EP-2011-117* (2011).

[61] The ATLAS Collaboration,*CERN-PH-EP-2011-114* (2011).

[62] The ATLAS Collaboration, *ATLAS-CONF-2011-046* (2011).

[63] The ATLAS Collaboration, *ATL-PHYS-PUB-2009-008* (2009).

[64] The ATLAS Collaboration, *ATLAS-CONF-2011-063* (2011).

[65] H. van der Graaf, H. Groenestege, C. Guyot, F. Linde, P. Rewiersma; *RasNiK, An Alignment System for the ATLAS MDT Barrel Muon Chambers, Technical System Description*, ETR 2000-04, NIKHEF Amsterdam (2001).

[66] http://lcgapp.cern.ch/project/persist/.

[67] http://root.cern.ch/drupal/.

[68] GEANT4 Collaboration, *Nucl. Instr. Meth* ***A506****: 250-303* (2003).

[69] K. Nakamura *et al.* (Particle Data Group), *J. Phys. G* ***37****, 075021* (2010) and 2011 partial update for the 2012 edition.

[70] The LHCb Collaboration *LHCb-CONF-2011-034* (2011).

[71] M. Morello, *Updated measur1e0ments of hadronic B decays at CDF"*, talk at EPS-HEP (2011), http://indico.in2p3.fr/contributionDisplay.py?contribId=1064&sessionId=2&confId=5116.

[72] G. Carlino "et. al., proceedings of "Final Workshop of the Grid Projects of the National Operation Program 2000-2006" edited by Consorzio COMETA, Italy, ISBN: 978-88-95-02-3 (2009).

[73] I.Bird et *al.*, *"LHC Computing Grid: Tecnichal Design Report"*

[74] E.Laure and B. Jones, *"Enabling Grids for e-Science: The EGEE Project"*, EGEE-PUB-2009-001, http://cdsweb.cern.ch/record/1128647, 2008.

[75] P.Eerola et *al.*, *"The NorduGrid architecture and tools"*, in Proceedings of CHEP 2004, eds. A.Aimar, J.Harvey and N.Knoors, CERN-2005-002,Vol.2, 2005.

[76] R.Pordes, *"The Open Science Grid (OSG)"* , in Computing in High Energy Physics and Nuclear Physics, http://cdsweb.cern.ch/record/865745, 2004.

[77] http://dashboard.cern.ch/atlas/.

[78] https://atlas-install.roma1.infn.it/atlas_install/.

[79] T.Sjostrand *et.al.*, *Comp. Phys. Commun.* ***135*** *238* (2001).

[80] M. Smizanska, *PythiaB an interface to Pythia6 dedicated to simulation of beauty events*, ATL-COM-PHYS-2003-038 (2003).

[81] http://root.cern.ch/root/html/TEfficiency.html.

[82] The ATLAS Collaboration, *ATLAS-CONF-2011-092* (2011).

[83] G.J. Feldman, R. Cousin, *Phys.Rev.D* ***57****:3873-3889* (1998).

[84] G. Punzi, Proceedings from *Phystat2003*, http://arxiv.org/abs/physics/0308063.

[85] W. A. Rolke,A. M. Lopez, J. Conrad, *Nucl.Instrum.Meth.* ***A551*** *493-503* (2004).

[86] A. L. Read, *J. Phys. G* ***28*** *2693*(2002).

[87] http://tmva.sourceforge.net/.